

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
6. September 2002 (06.09.2002)

PCT

(10) Internationale Veröffentlichungsnummer  
**WO 02/068682 A2**

- (51) Internationale Patentklassifikation?: C12Q 1/68, C12N 15/10, 15/39, C07K 14/07
- (21) Internationales Aktenzeichen: PCT/EP02/01909
- (22) Internationales Anmelde datum:  
22. Februar 2002 (22.02.2002)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität:  
101 08 626.1 22. Februar 2001 (22.02.2001) DE
- (71) Anmelder (*für alle Bestimmungsstaaten mit Ausnahme von US*): SAHIN, Ugur [TR/DE]; III Medizinische Klinik, Obere Zahlbacher Strasse 63, 55131 Mainz (DE).
- (71) Anmelder und  
(72) Erfinder: TUERECI, Özlem [DE/DE]; III Medizinische Klinik, Obere Zahlbacher Strasse 63, 55131 Mainz (DE). LUDEWIG, Burkhard [DE/DE]; III Medizinische Klinik, Obere Zahlbacher Strasse 63, 55131 Mainz (DE).
- (74) Anwälte: VOSSIUS, Volker usw.; Holbeinstrasse 5, 81679 München (DE).
- (81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Bestimmungsstaaten (regional): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

A2

(54) Title: METHOD FOR IDENTIFYING BIOLOGICALLY ACTIVE STRUCTURES OF MICROBIAL PATHOGENS

(54) Bezeichnung: VERFAHREN ZUR IDENTIFIZIERUNG BIOLOGISCH AKTIVER STRUKTUREN MIKROBIELLER ERREGER

(57) Abstract: The invention relates to a method for identifying biologically active structures which are coded by the genome by microbial pathogens, starting from genomic controlled nucleic acids.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft ein Verfahren zur Identifizierung biologisch aktiver Strukturen, die durch das Genom von mikrobiellen Erregern kodiert werden, ausgehend von genetischen Erregernukleinsäuren.

BEST AVAILABLE COPY

**Verfahren zur Identifizierung biologisch aktiver Strukturen mikrobieller Erreger**

- 5 Die vorliegende Erfindung betrifft ein Verfahren zur Identifizierung biologisch aktiver Strukturen, die durch das Genom von mikrobiellen Erregern kodiert werden, ausgehend von genetischen Erreger-nukleinsäuren.

**Hintergrund der Erfindung**

10

Die Entwicklung von molekular definierten Serodiagnostika und Impfstoffen setzt die molekulare Kenntnis und die Verfügbarkeit der durch das Immunsystem eines infizierten Wirtes erkannten Antigene des pathogenen Erregers (das mikrobielle Immunom) voraus. Die Serodiagnostik von Infektionskrankheiten basiert auf dem Nachweis von im Blut zirkulierenden Antikörpern, die spezifisch gegen immunogene Bestandteile (Antigene) des Erregers gerichtet sind und somit eine vorhandene oder abgelaufene Infektion anzeigen. Die Kenntnis dieser Antigene erlaubt es, dieselben als molekular definierte Impfstoffe rekombinant herzustellen. Diese Impfstoffe können einem Organismus einen Schutz vor einer Infektion mit dem betreffenden Erreger bieten (prophylaktische Immunisierung) aber auch bei persistierenden und chronischen Erregern zu deren Eliminierung dienen (therapeutische Immunisierung). Die Bedeutung, die solchen Antigenen sowohl für eine spezifische Diagnostik als auch für eine spezifische Therapie zukommt, führt zu einem beträchtlichen Interesse an der Identifikation dieser Strukturen.

Während für niedrigkomplexe Infektionserreger mit bekanntem Genom (z.B. einfache Viren) 25 die meisten relevanten Antigene molekular identifiziert sind, ist noch unklar, welche Antigene bei komplexeren Infektionserregern (z.B. Bakterien) immunologisch relevant sind. Die Ursache hierfür liegt darin, dass die komplexen Genome dieser Erreger eine Vielzahl von Genen (1000 bis über 4000) enthalten, die eine schnelle Identifikation der relevanten Antigene erschweren. Selbst bei Erregern, deren Genome vollständig bekannt sind, muss man 30 davon ausgehen, dass die Zuordnung aller Nukleotidbereiche zu tatsächlich für Proteine kodierenden -und somit potentiell Antigene ergebenden Abschnitten- noch nicht vollständig erfolgt ist.

Eine Reihe von Techniken, die zur Identifikation von Antigenen in den letzten Jahren 35 entwickelt wurden, versuchen dieser Komplexität gerecht zu werden. Die meisten dieser Verfahren stammen aus dem Feld der „Proteomics“-Technologien, welche Hochdurchsatz-Technologien der Proteinanalyse bezeichnet.

Eine dieser Hochdurchsatz-Technologien umfasst die Verwendung von 2-D-Gelen (z.B. Liu B, Marks JD. (2000), *Anal. Biochem.* 286,1191-28). Da bei diesem Verfahren große Mengen von Erregern benötigt werden, werden diese zunächst unter Kulturbedingungen angezüchtet. Danach werden Extrakte aus lysierten Erregern hergestellt und die darin befindlichen Proteine per Gelelektrophorese aufgetrennt. Durch Immunseren (Patientenserien, Seren immunisierter Tiere) erkannte Proteinkomplexe können durch Isolierung und Mikrosequenzierung analysiert werden. Dieses Verfahren weist eine Reihe von Limitationen und Nachteilen auf, da für die Untersuchungen große Mengen von Erregermaterial notwendig sind. Eine Analyse direkt aus der Primärläsion ist nicht möglich, sondern erst nach z.T. sehr langwieriger Kultur (z.B. bei Mykobakterien). Einige Erreger sind nicht einfach anzuzüchten, für unbekannte Erreger sind Kulturbedingungen oft nicht etabliert. Tote Erreger entziehen sich von vornehmlich diesem Anreicherungsverfahren.

Ein weiterer Nachteil der 2-D-Gel-Technologie besteht darin, dass sich der Genexpressionstatus eines Erregers in Zellkultur deutlich von dem *in vivo* unterscheidet.

Viele pathogene Genprodukte werden erst bei der Invasion des Erregers in den Wirtsorganismus eingeschaltet. Damit stehen für die Untersuchung lediglich solche Proteine zur Verfügung, die im Infektionserreger zum Zeitpunkt der Kultur exprimiert werden. Damit wird eine Reihe von Proteinen, die nur im Wirt unter Infektionsbedingungen in detektierbarer Menge exprimiert werden, von der Untersuchung ausgeschlossen. Gerade solche können jedoch für eine diagnostische Serologie, die eine klinisch nicht relevante Kolonialisierung von einer invasiven Infektion unterscheiden müssen, relevant sein.

Weiterhin werden mittels 2-D-Gelen Antigene als Proteine identifiziert. Die Nukleotidsequenz, die Grundlage für viele nachfolgende Analysen ist, muss erst noch ermittelt werden.

Als Alternative zum 2-D-Gel-Verfahren können bei Erregern, deren gesamte Nukleotidsequenz bekannt ist, alle putativen Gene in Expressionskassetten eingeführt, rekombinant exprimiert und auf Immunogenität oder Antigenität untersucht werden (Pizza et al. (2000), *Science* 287(5459):1816-20). Die rekombinant exprimierten Gene werden z.B. in einem parallelisierten Dot-Blot Verfahren mit Immunseren gescreent.

Die Nachteile dieses Verfahrens bestehen darin, dass lediglich Proteine, von denen bekannt ist, dass sie im interessierenden Erreger exprimiert werden, der Analyse zugänglich sind. Erreger mit unbekannter Nukleotidsequenz entziehen sich dieser Analyse.

Da die oben genannten Untersuchungstechnologien material-, zeit-, personal- und kostenaufwendig sind, sind sie nur wenigen großen Zentren vorbehalten.

Eine effiziente und potentiell wirksame Alternative zu den „Proteomics“-Ansätzen stellt das Immunoscreening genomicscher Expressionsbanken (z.B. Pereboeva et al. (2000), *J. Med. Virol.* 60: 144-151) dar. Allerdings müssen auch hierfür Infektionserreger unter definierten Kulturbedingungen angereichert werden. Das Genom der Erreger wird anschließend isoliert, enzymatisch oder mechanisch in Fragmente geschnitten und in Expressionsvektoren kloniert.

Die exprimierten Fragmente können dann mit Seren daraufhin überprüft werden, ob sie durch Seren von infizierten Organismen erkannt werden. Der Vorteil dieses Verfahrens ist, dass es kosteneffektiv und schnell zur Identifikation von Antigenen führen kann. Jedoch ist auch für dieses Verfahren aufgrund der großen Menge an benötigten Erregernukleinsäuren unumgänglich, dass die Erreger durch *in vitro* Kultur vermehrt und anschließend aufgereinigt werden. Damit ist das Verfahren bisher auf Erreger limitiert, für die Kultur- und Aufreinigungsmodalitäten bekannt und etabliert sind.

Ein Problem ist die Identifikation von bis dato nicht oder nur unzureichend charakterisierten Infektionserregern. Insbesondere [SH1] beschäftigt sich die vorliegende Erfindung mit den folgenden Aufgabenstellungen:

- 1) Entzündliche Erkrankungen mit aufgrund von Epidemiologie und klinischem Verlauf als wahrscheinlich anzunehmender infektiöser Ursache für die bisher mit bekannten Verfahren Erreger nicht definiert bzw. nur unzureichend charakterisiert werden können. Hierzu zählen Erkrankungen wie Multiple Sklerose, Kawasaki-Disease, Sarkoidose, Diabetes mellitus, Morbus Whipple, Pityriasis rosea u.a. Wünschenswert für diese Erkrankungen ist ein Verfahren, welches eine systematische Analyse auf bisher unbekannte Infektionserreger von primären Patientenmaterial z.B. Lymphknoten-Biopsien erlaubt.
  - 2) Newly emerging infectious diseases. Hierzu zählen Infektionskrankheiten, die durch bisher nicht bekannte (z.B. HIV in den 80er Jahren) oder nicht gut charakterisierte Krankheitserreger verursacht werden und aufgrund z.B. Änderung der Epidemiologie plötzlich im Focus des klinischen Interesses stehen. Medizinisch und sozioökonomisch essentiell für diese Klasse von Infektionserkrankungen ist das schnell Erreger identifiziert und entsprechende Diagnostika und eventuell auch Impfstoffe produziert werden können. Da im allgemeinen die Etablierung der Kultivierungsbedingungen für nicht charakterisierte Erreger bis zu mehrere Jahre in Anspruch nehmen kann, ist auch in diesem Fall eine Identifikation von Erregern und Erregerantigenen direkt aus dem infizierten Gewebe sehr wünschenswert, kann aber durch bekannte Verfahren nicht geleistet werden.
- Es besteht daher ein großer Bedarf an einem Verfahren, welches die direkte Verwendbarkeit von Primärmaterial zur Erregeridentifikation ohne Notwendigkeit einer Erregervermehrung durch Kultur erlaubt. Weiterhin sollte dieses Verfahren die effiziente Entdeckung bisher unbekannter Krankheitserreger in Primärmaterial ermöglichen.

**Zusammenfassende Beschreibung der Erfindung**

Eine Aufgabe der vorliegenden Erfindung war es daher ein Verfahren zu entwickeln, welches eine Identifikation von Erregernukleinsäuren direkt aus einer limitierten Menge (z.B. 50 mm<sup>3</sup>) infizierten Patientenmaterial ermöglicht.

- 5 Die vorliegende Erfindung beschreibt ein Verfahren zur systematischen Identifikation von bekannten wie auch unbekannten nukleinsäurekodierten Erregern und ihren Antigenen anhand der von ihnen ausgelösten Immunantwort im Wirtsorganismus.

Gegenstand der vorliegenden Erfindung ist daher ein Verfahren zur Identifizierung von biologisch aktiven Strukturen, die durch das Genom von mikrobiellen Erregern kodiert

- 10 werden, ausgehend von genomischen Erregernukleinsäuren, wobei das Verfahren die folgenden Schritte umfasst: Gewinnung von genomischen Erregernukleinsäuren aus erregerhaltigen Proben, sequenzunabhängige Amplifikation von genomischen Erregernukleinsäuren, Expression amplifizierter Erregernukleinsäuren und Screening und Identifizierung der biologisch aktiven Struktur.

- 15 Das erfindungsgemäße Verfahren hat den entscheidenden Vorteil, dass auch bei sehr geringer Erregermenge eine umfassende Identifikation vom Wirtsorganismus erkannter Erregerantigene (mikrobielles Immunom) möglich ist.

- Das erfindungsgemäße Verfahren ist dadurch gekennzeichnet, dass geringste Ausgangsmengen bis minimal 1 pg Erregernukleinsäure ausreichen, um eine effektive 20 Analyse zu erzielen. In einer bevorzugten Ausführungsform werden 10-20 pg, insbesondere 1-10 pg an Erregernukleinsäure verwendet.

- Die hohe Sensitivität des Verfahrens ermöglicht einerseits die Untersuchung von Erregern aus Primärisolaten ohne die Notwendigkeit, diese Erreger zuvor durch *in vitro* Kultur anreichern zu müssen. Somit können Erreger untersucht werden, die sich mit bekannten Methoden 25 schwer (z.B. Mykobakterium tuberculosis u.a.) oder gar nicht anzüchten lassen (z.B. Mykobacterium leprae oder nicht-vitale Keime).

- Andererseits wird durch den Wegfall der *in vitro* Anreicherung vermieden, dass es zu einer durch die *in vitro* Kultur bedingten Verfälschung der Keimpopulation kommt (z.B. durch Überwachsen relevanter pathogener Keime bei Mischnfektionen).

- 30 Weiterhin eröffnet die hohe Sensitivität des Verfahrens die Möglichkeit, ein breites Spektrum an Ausgangsmaterialien zur Erregerisolierung zu nutzen. Der Begriff „Proben“ bezeichnet dabei verschiedene biologische Materialien wie Zellen, Gewebe, Körperflüssigkeiten. In einer bevorzugten Ausführungsform der Erfindung stellen Blut, Gewebe, kultivierte Zellen, Serum, Sekret aus Läsionen (Pusteln, Krusten etc.) und andere Körperflüssigkeiten wie z.B. Urin, 35 Speichel, Liquor, Gelenkflüssigkeiten, Galle und Augendrüsensekrete bevorzugt verwendete Proben dar.

## Detaillierte Beschreibung der Erfindung

### **Gewinnung von genomischen Erregernukleinsäuren aus erregerhaltigen Proben**

Im ersten Schritt des erfindungsgemäßen Verfahrens werden genomicsche

5 Erregernukleinsäuren aus erregerhaltigen Proben gewonnen.

Der hier verwendete Begriff „mikrobieller Erreger“ umfasst virale und bakterielle Erreger. Erreger liegen in Wirtszellen oder im Zellverband mit Wirtszellen vor und müssen mit Ausnahme von frei im Serum zirkulierenden Erregern zugänglich gemacht werden. In einer bevorzugten Ausführungsform der Erfindung liegt der Erreger intrazellulär oder extrazellulär

10 vor.

Intrazelluläre Erreger können durch Zellyse (z.B. mechanisch oder über Detergenzien, bei eukaryontischen Zellmembranen z.B. mittels SDS) freigesetzt werden. Bevorzugte SDS-Konzentration sind für gram positive Bakterien >0,05% bis 1% und für gramnegative Bakterien 0,05% bis 0,1%. Der Fachmann kann die geeignete Konzentration für andere

15 Detergenzien leicht bestimmen, indem die Konzentration eingesetzt wird bei der die Hülle, z.B. die Bakterienwand des grampositiven oder des grammnegativen Bakteriums noch intakt ist, die eukaryontische Zellwand aber bereits aufgelöst wird.

Extrazelluläre Erreger können z.B. über ihre deutlich geringere Partikelgröße (20 nm-1 µM) von Wirtszellen getrennt werden (z.B. durch Sedimentation/Zentrifugation und/oder Filtration).

In einer besonders bevorzugten Ausführungsform der Erfindung ist der Erreger nicht-infektiös und/oder nicht-vital. Die hohe Sensitivität des Verfahrens ermöglicht es, dass auch auf nicht-infektiöse und/oder nach Ablauf einer floriden Infektion auf residual verbliebene nicht-vitale Erreger zurückgegriffen werden kann.

25 Im folgenden wird der erste Schritt des erfindungsgemäßen Verfahrens, die Gewinnung genomicsche Erregernukleinsäuren aus erregerhaltigen Proben weiter beschrieben.

Die Gewinnung genomicscher Erregernukleinsäuren aus erregerhaltigen Proben ist in einer bevorzugten Ausführungsform der Erfindung durch die folgenden Schritte gekennzeichnet: Freisetzung von Erregerpartikeln aus erregerhaltigen Proben, nachfolgend Elimination und/oder Reduktion kontaminierender Wirtsnukleinsäuren und abschließend Extraktion der genomicschen Erregernukleinsäure aus freigesetzten Erregerpartikeln.

30 Gemäss einer Ausführungsform der Erfindung erfolgt die Freisetzung von Erregerpartikeln durch Zellyse, Sedimentation, Zentrifugation und/oder Filtration.

35 Gemäss einer bevorzugten Ausführungsform erfolgt die Elimination der kontaminierenden Wirtsnukleinsäuren dadurch, dass vor der Extraktion der Erregernukleinsäure ein RNase- und/oder DNase-Verdau durchgeführt wird.

Dies ist besonders dann vorteilhaft, wenn der Erreger ein durch Kapsidproteine geschütztes Virus ist. Da die Kapsidhülle von Viren einen Schutz vor Nukleaseen bietet, sind die Erreger entsprechend vor der Aktivität extrazellulärer RNasen und DNasen geschützt.

- Eine weitere Ausführungsform der Erfindung umfasst daher den Schritt der Elimination und/oder Reduktion kontaminierender Wirtsnukleinsäuren durch RNase und/oder DNase-Verdau besonders für virale Erreger (s. Beispiel Vaccinia-virus). Die Dnase Behandlung zur Aufreinigung von Viruspartikeln ist dem Fachmann bekannt und kann durchgeführt werden wie beispielsweise beschrieben bei Dahl R, Kates JR., Virology 1970 ;42(2):453-62, Gutteridge WE, Cover B., Trans R Soc Trop Med Hyg 1973;67(2):254), Keel JA, Finnerty WR, Feeley JC., Ann Intern Med 1979 Apr;90(4):652-5 oder Rotten S. in Methods in Mycoplasmology, Vol. 1, Academic Press, 1983.

Eine weitere Möglichkeit, Erreger wie grampositive oder gramnegative Bakterien aus Geweben anzureichern, ist wie im erfindungsgemäßen Verfahren angewandt, eine differenzielle Lyse (nachfolgend auch als sequentielle Lyse bezeichnet) infizierter Gewebe mit Detergenzien wie z.B. SDS, die Lipidmembranen auflösen. Hierbei macht man sich den Umstand zu Nutze, dass die Zellmembranen eukaryontischer Wirtszellen sehr viel empfindlicher auf niedrige Konzentration von SDS reagieren und aufgelöst werden, während Bakterienwände widerstandsfähiger sind und ihre korpuskuläre Integrität erhalten bleibt.

- Nach Anreicherung werden die Erregernukleinsäuren von den korpuskulären Bestandteilen des Erregers abgetrennt und aus dem Erreger freigesetzt. Dies kann mittels dem Fachmann bekannten Standardtechniken durchgeführt werden, wobei in einer bevorzugten Ausführungsform der Erfindung die Abtrennung durch Proteinase K-Verdau, Denaturierung, Hitze- und/oder Ultraschallbehandlung, enzymatisch mit Lysozym-Behandlung oder organische Extraktion erfolgt.

Die Aufreinigung von Bakterien/Viren durch Detergentien wie SDS kann durchgeführt werden wie bei Takumi K, Kinouchi T, Kawata T., Microbiol Immunol 1980;24(6):469-77, Kramer VC, Calabrese DM, Nickerson KW., Environ Microbiol 1980 40(5):973-6, Rudoi NM, Zelenskaia AB, Erkovskaia., Lab Delo 1975;(8):487-9 oder Keel JA, Finnerty WR, Feeley JC., Ann Intern Med 1979 Apr;90(4):652-5, s.a. US Patent 4,283,490.

Die Erfindung löst das Problem, dass in einem infizierten Gewebe / Organ nur ein Teil des Gewebes/Zellen mit dem Erreger infiziert ist.

- Die vorliegende Erfindung ist auch geeignet zum Nachweis von Erregern bei Fällen wie z.B. Mykobakteriosen, wo nur wenige bis einzelne Erregerpartikel in den infizierten Geweben existent sind.

Die vorliegende Erfindung hat den Vorteil, das in den Fällen einer Infektion mit nur einer geringen Totalmenge an Erregernukleinsäuren pro Gewebeinheit (z.B. 50 mm<sup>3</sup>) der Nachweis der Erregers möglich ist. Als Zahlenbeispiel für DNA-kodierte Pathogene lässt sich errechnen das ca. 10<sup>6</sup> Viren mit einer durchschnittlichen Genomgrösse von 10000 Basen bzw. 5 10<sup>4</sup> Bakterien mit einer durchschnittlichen Genomgröße von 1.000.000 Basen gerade 10 pg Erregernukleinsäuren beinhalten. Zudem ist das Genom von den meisten Infektionserregern deutlich kleiner als das menschliche Genom (bis Faktor 10<sup>6</sup> bei Viren, bis Faktor 10<sup>4</sup> bei Bakterien). Dadurch verdünnt sich der Anteil der Gesamtmenge Erregernukleinsäuren gegenüber der Gesamtmenge der Wirtsnukleinsäuren je nach Kopienzahl und Genomgrösse 10 des Erregers um eine vielfache Zehnerpotenz (Verhältnis Wirtsnukleinäure / Erregernukleinsäure 10<sup>4</sup> bis 10<sup>9</sup>).

### Sequenzunabhängige Amplifikation von genomischen Erregernukleinsäuren

Das erfindungsgemäße Verfahren umfasst weiterhin die sequenzunabhängige Amplifikation 15 genomischer Erregernukleinsäuren mittels Polymerase-Kettenreaktion (PCR), die der Vermehrung des Ausgangsmaterials dient.

Hierbei werden PCR-Primer mit Zufallssequenzen (Random-Oligonukleotide) verwendet, so dass nicht nur spezifische Genbereiche, sondern möglichst das gesamte Genom des Erregers repräsentativ amplifiziert werden kann (unselektive Nukleinsäureamplifikation mit 20 degenerierten Oligonukleotiden). Diese Primer binden natürlich auch an Wirts-DNA. Diese wurden jedoch, wie oben beschrieben, durch vorhergehenden RNase- und/oder DNase- Verdau bzw. durch selektive Zellyse reduziert bzw. eliminiert.

Allerdings ist Eliminierung bzw. Reduktion von Wirts-DNA durch vorhergehenden RNase- und/oder DNase-Verdau bzw. durch selektive Zellyse nicht zwingend erforderlich in dem 25 erfindungsgemässen Verfahren. Bei einer hohen Konzentration der Erreger- DNA oder RNA, wie sie zum Beispiel in Pusteln oder Virusbläschen ist eine Eliminierung bzw. Reduktion von Wirts-DNA nicht erforderlich.

Der hier verwendete Begriff „genomische Erregernukleinsäure“ umfasst sowohl genomische DNA als auch RNA.

30 In einer bevorzugten Ausführungsform der Erfindung ist die genomische Erregernukleinsäure DNA, und ihre sequenzunabhängige Amplifikation erfolgt durch Klenowreaktion mit Adaptoroligonukleotiden mit degeneriertem 3'-Ende und nachfolgender PCR mit Oligonukleotiden, die der Adaptorsequenz entsprechen.

In einer weiteren bevorzugten Ausführungsform ist die genomische Erregernukleinsäure 35 RNA, und ihre Amplifikation erfolgt durch reverse Transkription mit degenerierten Oligonukleotiden und nachfolgender Amplifikation durch PCR .

Wenn nicht bekannt ist, ob der Erreger durch RNA oder DNA kodiert wird, werden beide Reaktionen in getrennten Reaktionsgefäßern angesetzt und getrennt amplifiziert. Die Amplikons repräsentieren in beiden Fällen genomische Nukleinsäurefragmente. Durch die Verwendung von degenerierten Oligonukleotiden wird eine zufällige Amplifikation (Shotguntechnik) des Gesamtgenoms ermöglicht.

Die Stärke des erfindungsgemäßen Verfahrens liegt in seiner hohen Sensitivität und Effizienz (Ausgangsmengen von wenigen picogramm reichen aus) bei gleichzeitig guter Erhaltung der Repräsentanz sämtlicher Bereiche des Gesamtgenoms. Die gute Repräsentanz der mikrobiellen Genabschnitte in den nach dem erfindungsgemäßen Verfahren generierten Bibliotheken wird durch Variationen in der Zwei-Schritt-PCR, wie z.B. Veränderungen der Salzkonzentration, erreicht. Die hohe Effizienz der Methode mit weitgehender Erhaltung der Repräsentanz erübriggt daher das Anlegen einer Primärkultur zur Vermehrung des Erregers mit allen durch sie bedingten Limitationen. Im Falle von unbekannten Erregern sind Kulturbedingungen nicht definiert und müssten im trial-and-error-Verfahren approximiert werden. Auch ist eine Reihe bekannter Erreger schwierig anzuzüchten. Bei Mischinfektionen können Primärkulturen durch Überwachungsphänomene gerade die relevante Erregerpopulation ausdünnen. Tote Erreger würden gar nicht erfasst. Diese Nachteile anderer Techniken werden durch das erfindungsgemäße Verfahren umgangen.

20

In einer bevorzugten Ausführungsform wird die Sequenz-unspezifische Amplifikation von Erreger-nukleinsäuren in zwei sequentiell hintereinandergeschalteten PCR-Schritten von jeweils 35-40 Zyklen durchgeführt.

Hierzu wird nach der ersten Amplifikation 1/20 bis 1/50 des Volumens der ersten PCR für die Reamplifikation unter variierenden Bedingungen (z.B. Variation der MgCl Konzentration, der Pufferbedingungen, der Polymerasen) genutzt. Durch die Reamplifikation in einer zweiten PCR wird wie im Beispiel 3A dargestellt, eine höhere Sensitivität des Verfahrens gewährleistet. Die Variation der Re-Amplifikationsbedingungen ermöglicht (siehe Beispiel 6, Abbildung 6) eine besonders gute Repräsentation unterschiedlicher Abschnitte des Erregergenoms und damit eine umfassende Analyse des Erregers.

#### Expression amplifizierter Erreger-nukleinsäuren

Auf die Amplifikation der genomischen Erreger-nukleinsäuren folgt deren Expression. Dazu werden die Erreger-nukleinsäuren zur Erstellung einer genomischen Expressionsbank des Erregers in geeignete Expressionsvektoren kloniert.

In einer Ausführungsform der Erfindung sind die Expressionsvektoren ausgewählt aus der Gruppe der viralen, eukaryotischen oder prokaryotischen Vektoren. Im Rahmen der Erfindung können alle Systeme genutzt werden, die eine Expression von rekombinanten Proteinen erlauben.

- 5 Nach dem Einbringen der Erregernukleinsäuren in die Vektoren werden die Vektoren bevorzugt in Lambda-Phagen verpackt.

In einer bevorzugten Ausführungsform der Erfindung wird die Expression der Erregernukleinsäuren durch Einbringen der Erregernukleinsäure in Lambda-Phagen Vektoren (z.B. Expressionsvektor Lambda-ZAP-Express, US Patent Nr. 5.128.256) gewährleistet.

- 10 Alternativ können auch andere Vektoren, die dem Fachmann bekannt sind, eingesetzt werden, besonders bevorzugt sind filamentöse Phagen-Vektoren, eukaryontische Vektoren, retrovirale Vektoren, adenovirale Vektoren oder Alphavirusvektoren.

15 **Screening und Identifizierung der biologisch aktiven Struktur.**

Der abschließende Schritt des erfindungsgemäßen Verfahrens umfasst das Screening der genomischen Expressionsbank und die Identifizierung der biologisch aktiven Struktur des Erregers durch Immunantworten infizierter Wirte.

- 20 Der hier verwendete Begriff „biologisch aktive Struktur“ bezeichnet Erregerantigene, enzymatisch aktive Proteine oder Pathogenitätsfaktoren des mikrobiellen Erregers.

Gemäß einer bevorzugten Ausführungsform der vorliegenden Erfindung stellt das Screening ein Immuno screening auf Erregerantigene dar und die Identifikation von Erregerantigenen umfasst die folgenden Schritte: Infektion von Bakterien mit Lambda-Phagen, Kultivierung der infizierten Bakterien unter Bildung von Phagenplaques, Transfer der Phagenplaques auf

- 25 eine Nitrozellulosemembran (oder einer anderen Festphase, die zur Immobilisierung von rekombinanten aus den Erregern abgeleiteten Proteinen geeignet ist), Inkubation der Membran mit Serum oder antikörperhaltigen Körperflüssigkeiten des infizierten Wirtes, Waschen der Membran, Inkubation der Membran mit sekundärem alkalischer Phosphatase-gekoppeltem anti-IgG-Antikörper der spezifisch für Immunglobuline des infizierten Wirtes  
30 ist, Nachweis der mit Wirts serum reaktiven Klone durch Farbreaktion, und die Isolierung und Sequenzierung der reaktiven Klone.

- Grundsätzlich ist es zur Identifikation der biologischen aktiven Struktur auch möglich, Erregernukleinsäuren in rekombinante filamentöse Phagen-Vektoren (z.B. pJufo) einzubringen, die eine Expression von Antigenen direkt auf den Oberflächen der filamentösen Phagen erlauben. In diesem Fall würde die Identifikation von Erregerantigenen die folgenden Schritte umfassen: Generierung von rekombinanten filamentösen Phagen durch Einbringen

der filamentösen Phagenvektoren in Bakterien, Inkubation generierter rekombinanter filamentöser Phagen mit Serum eines infizierten Wirtes, Selektion der filamentösen Phagen, an die Immunglobuline des Wirts gebunden haben, über immobilisierte Reagenzien die spezifisch für die Immunglobuline des infizierten Wirts sind, und Isolierung 5 und Sequenzierung der selektierten Klone.

Vom Erregergenom abgeleitete und rekombinant exprimierte Proteine können z.B. auf Festphase gebunden oder im Rahmen eines Panning-/Captureverfahrens mit spezifischen Immunantwortäquivalenten des infizierten Wirts gescreent werden. Dies sind zum einen 10 Antikörper verschiedener Immunglobulinklassen-/subklassen, in erster Linie IgG. Zu diesem Zweck wird Wirtsserum verwendet. Dies sind aber auch spezifische T-Lymphozyten gegen MHC-restringiert erkannte Epitope von Erregerantigenen, die in einem eukaryotischen System getestet werden müssen.

Die Herstellungsbedingungen für die genomische Bank erfolgt so, dass die inserierten 15 Fragmente aufgrund der Eigenart der PCR-Primer nach dem Zufallsgeneratorprinzip entstehen. Entsprechend sind Bereiche aus bekannten Antigenen repräsentiert, die natürlicherweise auch als Proteine gebildet würden. Auch Fragmente aus normalerweise nicht exprimierten intergenen Bereichen können entstehen, die je nach Länge offener Leseraster zur Expression kurzer Nonsenseproteine oder -peptide führen können. Ein wichtiger Aspekt ist, 20 dass im erfundungsgemäßen Verfahren automatisch auch bisher nicht identifizierte Erregerproteine vorliegen können.

#### Detalierte Beschreibung bevorzugter Ausführungsformen der vorliegenden Erfindung

Die vorliegende Erfindung kombiniert die Expression der Gesamtdiversität aller erdenklichen 25 rekombinanten Proteine mit der anschließenden Anwendung eines sehr stringenten Filters, nämlich die in infizierten Wirten im Rahmen des natürlichen Verlaufes der Erkrankung entstehende spezifische Immunantwort.

In einer weiteren Ausführungsform der Erfindung werden die Erreger vor der Nukleinsäureamplifikation durch Fällung mit Polyethylenglykol, Ultrazentrifugation, Gradientenzentrifugation oder durch Affinitätschromatographie angereichert. Dieser Schritt 30 ist jedoch nicht zwingend vorgesehen.

Die Fällung mit Polyethylenglykol ist vor allem bei viralen Partikeln effizient. Alternativ bietet sich bei bekannten Erregern die Affinitätschromatographie unter Nutzung von 35 erregerspezifischen Antikörpern gegen definierte und stabile Oberflächenstrukturen an. Eine weitere Alternative ist bei unbekannten Erregern die Nutzung des polyklonalen Patientenserums selbst möglich, wobei das polyklonale Patientenserum auf der Festphase immobilisiert und zur Affinitätsanreicherung von Erregern als spezifisches Capture-Reagenz verwendet wird.

Das hier beschriebene Verfahren kann als Plattformtechnologie eingesetzt werden, um hocheffizient Antigen-kodierende Erreger-nukleinsäuren aus geringsten Mengen erregerhaltigen Materials zu identifizieren. Wie in den nachfolgenden Beispielen gezeigt, sind 1 bis 20 pg genomische Nukleinsäure ausreichend, um eine umfassende Identifikation des 5 durch natürliche Immunantworten serologisch erkannten Antigenrepertoires einzelner Erreger zu leisten. So konnten mit der hier beschriebenen Technologie z.B. für das Vaccinia-virus als immunodominant bekannte Antigene ausgehend von 20 pg DNA identifiziert werden.

Die geringe Menge an Nukleinsäuren, die für dieses Verfahren benötigt wird, ermöglicht die 10 Anwendung des Verfahrens für medizinisch wichtige Fragestellungen, die mit den bisher bekannten Verfahren schlecht oder gar nicht bearbeitet werden konnten.

Hierzu zählt z.B. die systematische direkte Identifikation von Erreger-nukleinsäuren aus infizierten Zellen, z.B. empfängliche *in vitro* Zelllinien, Organe, entzündliche Läsionen wie z.B. Pusteln auf der Haut, auf Schleimhäuten, aus infizierten inneren lymphatischen und nicht-lymphatischen Organen oder aus erregerhaltigen Flüssigkeiten (z.B. Speichel, Sputum, 15 Blut, Urin, Eiter, Ergüsse), die von infizierten Organismen gewonnen werden. Ausgehend von einer Sensitivität von 1 pg sind je nach Genomgrösse des Erregers (z.B. Viren 3.000 - 250.000 bp oder Bakterien 100.000 - 5.000.000 bp) 50 bis  $10^5$  Erregerpartikel ausreichend, um für Antigene kodierende Erreger-nukleinsäuren zu identifizieren. Es ist zu erwarten, dass 20 in den meisten Fällen die Anzahl von Erregerpartikeln weit über der Sensitivitätsgrenze des erfindungsgemäßen Verfahrens liegt.

Die hohe Sensitivität des erfindungsgemäßen Verfahrens erlaubt, wie bereits vorstehend beschrieben, eine Untersuchung von Erregern aus Primärisolaten ohne Notwendigkeit einer *in vitro* Kultur. Besonders wichtig ist, dass sehr kleine Mengen von Ausgangsmaterial, wie z.B. Stecknadelkopf-große Biopsien oder wenige Microliter infizierte Proben-Flüssigkeiten für 25 die erfolgreiche Anreicherung und Identifikation von biologisch aktiven Strukturen durch das erfindungsgemäße Verfahren ausreichen. Hiermit kann das erfindungsgemäße Verfahren auf jegliches Überschussmaterial aus der medizinisch-klinischen Diagnostik und auf kryoarchivierte Probenmaterialien zugreifen (im Beispiel 9 dargestellt).

Die Sequenzierung der anhand des erfindungsgemäßen Verfahrens identifizierten 30 Erreger-nukleinsäure führt zur Identifikation des Erregers, aus dem die Nukleinsäure ursprünglich stammt. Damit ist eine vorherige Kenntnis des Erregers nicht notwendig und das Verfahren eignet sich zur Entdeckung von bisher nicht definierten Infektionserregern bzw. zu bisher nicht bekannten Antigenen.

Das Verfahren kann somit zur Untersuchung einer Reihe von Erkrankungen eingesetzt 35 werden, bei denen ein Infektionserreger ätiologisch vermutet wird, aber noch nicht identifiziert werden konnte. Hierzu zählen Erkrankungen, welche die Koch'schen Postulate teilweise erfüllen (z.B. Übertragbarkeit), bei denen jedoch die Keime aufgrund von fehlender Anzüchtbarkeit/Isolierbarkeit der Erreger nicht identifiziert werden können. Weitere Beispiele

sind Erkrankungen wie die Sarkoidose, Pitryasis rosea, Multiple Sklerose, Diabetis Mellitus und Morbus Crohn.

Auch bei einem Teil von Patienten mit ätiologisch unklarer chronischer Hepatitis (chronische non-B, non-C Hepatitis) wird eine bisher unbekannte Viruserkrankung vermutet. Die

- 5 Keimzahlen im Serum von Patienten mit bekannter chronischer Virushepatitis sind hoch. So finden sich bei Patienten mit infektiöser chronischer HBV- bzw. HCV-induzierter Hepatitis in 1 ml Blut  $10^7$ - $10^8$  Hepatitis B- bzw. Hepatitis C-Partikel. Unter der Annahme einer annähernd vergleichbaren Keimzahl ist das erfundungsgemäße Verfahren auch dazu geeignet, putative non-B, non-C Hepatitis-Erreger ausgehend aus geringen Mengen Blut/Serum (1-10 ml)
- 10 infizierter Patienten zu identifizieren.

Der Schritt des Immunoscreening der vorliegenden Erfindung als hochsensitives und hochspezifisches Hochdurchsatz-Detektionsverfahren erlaubt die Identifikation einer Antigen-kodierenden Nukleinsäure unter  $10^6$ - $10^7$  nicht-immunogenen Klonen.

- 15 Hierzu reicht eine geringe Reinheit der Erreger-nukleinsäuren zur Identifizierung des Erregers aus. Diese geringe Reinheit kann problemlos durch verschiedene aus dem Stand der Technik bekannte Verfahren, wie z.B. Präzipitation von Erregerpartikeln mit Polyethylenglykol (PEG) und/oder Affinitätschromatographie und/oder Degradation von kontaminierenden Wirtsnukleinsäuren mit Nukleaseen erreicht werden.
- 20 Eine zusätzliche Möglichkeit zur Anreicherung von Erregerpartikeln ist die Nutzung der in infizierten Organismen gegen Erregerpartikel gebildeten spezifischen Antikörper für Capture-Verfahren.

Das Immunoscreening als integraler Bestandteil des Verfahrens erlaubt die Analyse von  $10^6$ - $5 \times 10^6$  Klonen innerhalb von kurzer Zeit (2 Monaten) durch eine einzige Person. Die

25 Kombination aus sequenzunabhängiger Amplifikation und serologischer Untersuchung mit hohem Durchsatz aller Nukleinsäureabschnitte in allen 6 Leserastern erlaubt schon bei mäßiger Reinheit der Ausgangsnukleinsäuren (Erreger-nukleinsäuren > 1% der Gesamt-nukleinsäuren) die umfassende Untersuchung aller potentiell für Polypeptide kodierenden Regionen unabhängig vom aktuellen Expressionstatus. Durch die Untersuchung

30 von genomischen Nukleinsäuren werden alle Genregionen erfasst, auch Gene, die nur zu bestimmten Zeitpunkten (z.B. nur in bestimmten Infektionszeitphasen) eingeschaltet werden. Dies erlaubt nicht nur eine Aussage über einzelne Antigene, sondern gibt auch Informationen über die Gesamtheit der Nukleotid-kodierten immunogenen Regionen (Immunom, siehe Abbildung 4A-C und 5). Über die Identifikation multipler z.T. überlappender Fragmente wird

35 darüberhinaus eine Einengung des serologisch erkannten Epitops innerhalb eines identifizierten Antigens ermöglicht (siehe Abbildung 4A-C). Die Stärke des Signals ermöglicht eine weitere Diskriminierung von dominanten und nichtdominanten Epitopen (siehe Abbildung 4A-C). Die identifizierten Erreger-Nukleinsäurefragmente sind direkt für die Entwicklung von Serodiagnostika und Impfstoffen verfügbar. Die identifizierte

Nukleinsäure kann als Matrize für die Entwicklung hochsensitiver direkter Erreger-nachweismethoden z.B. durch Anwendung von Nukleinsäure-spezifischer Amplifikation über die Polymerasekettenreaktion (PCR) genutzt werden. Die identifizierten Fragmente können auch für die Entwicklung von Diagnosetests basierend auf dem Nachweis von Antigen-spezifischen T-Lymphozytenreaktionen genutzt werden.

Die Abgrenzung gegen Technologien wie „Proteomics“ ist schon in der Einleitung diskutiert worden. Das erfindungsgemäße Verfahren unterscheidet sich technisch gegenüber zwei weiteren verwandten Methoden: die serologische Untersuchung von genomischen Erregerbibliotheken und der SEREX-Technologie.

- 10 Für die serologische Untersuchung von genomischen Bibliotheken wurden von einigen Gruppen (z.B. Luchini et al. (1983), *Curr Genet* 10:245-52, Bannantine et al. (1998), *Molecular Microbiology* 28: 1017-1026) Expressionsbibliotheken aus aufgereinigter mechanisch zerkleinerter oder enzymatisch verdauter Erreger-DNA hergestellt. Für die Herstellung von Expressionsbibliotheken mit diesem Verfahren werden zur Herstellung repräsentativer Banken je nach Größe des Genoms zwischen 0,5 – 5 µg an aufgereinigten Erreger-nukleinsäuren benötigt (Faktor 10<sup>5</sup>-10<sup>6</sup> Mehrbedarf an Erreger-nukleinsäuren). Hiermit verschließt sich die Methode der Untersuchung von Erregern aus Primärisolaten, in denen weitaus geringere Mengen an Erregern und Erreger-nukleinsäuren vorhanden sind. Es ist hier unumgänglich, dass die Erreger isoliert, *in vitro* angezüchtet und anschließend aufwendig aufgereinigt werden. Damit müssen als Grundvoraussetzungen zur Anwendung dieser Methode die Kultur- und Aufreinigungsmodalitäten für den jeweiligen Erreger bekannt und zuvor etabliert sein. Dies ist jedoch gerade für viele Viren und intrazelluläre Erreger technisch aufwendig und verlangt eine große Expertise. Damit entfällt auch die Möglichkeit, unbekannte Erreger, sowie solche, die nicht mehr vital sind, zu identifizieren.
- 15
- 20
- 25 Eine weitere Abgrenzung ist gegenüber der mit SEREX (Sahin et al. (1995), *Proc Natl Acad Sci USA* 92: 11810-3; Sahin et al. (1997), *Curr Opin Immunol* 9, 709-716) bezeichneten Methode vorzunehmen. Für SEREX wird aus erkrankten Geweben mRNA extrahiert, cDNA-Expressionsbibliotheken hergestellt und mit Seren desselben Individuums, aus dem das Gewebe stammt, auf immunreaktive Antigene gescreent.
- 30 Ein wesentlicher Unterschied zum erfindungsgemäßen Verfahren besteht darin, dass für die SEREX-Methode cDNA-Expressionsbibliotheken aus Wirtszellen infizierter Gewebe verwendet werden.

Die unterschiedliche Qualität und der unterschiedliche Ursprung der Nukleinsäuren führt zu folgenden Unterschieden:

- 35 Die Verwendung von Gesamt-mRNA aus Wirtszellen erhöht bei SEREX die Komplexität der Bibliothek und vermindert die Wahrscheinlichkeit der Identifikation Erreger-abgeleiteter Transkripte. Für tierische Wirtszellen ist von 40.000 bis 100.000 unterschiedlichen wirtseigenen Transkripten auszugehen. Die Anzahl der Transkripte für die meisten Erreger liegt weitaus geringer (Viren 3-200 Transkripte, Bakterien 500-4000 potentielle

Genprodukte). Zusätzlich ist in den meisten Fällen nur ein geringer Teil von Wirtszellen mit Erregern infiziert, so dass der Anteil Erreger-abgeleiteter Nukleinsäuren in der Gesamt-mRNA Population sich weiter verdünnt. Da viele wirtseigene Transkripte auch für natürliche oder krankheitsassoziierte Autoantigene kodieren (Sahin et al., 2000, Scanlan et al. (1998), *Int J Cancer* 76, 652-658) ist die Identifikation von Erreger-abgeleiteten Antigenen durch die präferentielle Detektion von Wirts-Gewebsautoantigenen sehr erschwert. Dieser Umstand ist verantwortlich dafür, das mit der SEREX-Technologie bei der Untersuchung von mehreren infizierten Geweben, wie z.B. von HBV-Ag+ Leberzell-Karzinomen (Scanlan et al. (1998), *Int J Cancer* 76, 652-658; Stenner et al. (2000), *Cancer Epidemiol Biomarkers Prev* 9, 285-10 90) bisher noch keine Erregerantigene identifiziert worden sind.

Aus den nachfolgenden Abbildungen und Beispielen wird ersichtlich, dass mit dem Verfahren der vorliegenden Erfindung aus geringsten Mengen von Erreger-nukleinsäuren immunologisch relevante virale und bakterielle Antigene identifiziert und charakterisiert werden können. Die in den nachfolgenden Beispielen identifizierten viralen Antigene waren dabei über das gesamte Genom des Vaccinia-virus verteilt, was auf eine gute Repräsentanz der verschiedenen Gene in der mit Hilfe des erfindungsgemäßen Verfahrens amplifizierten DNA schließen lässt (siehe Abb. 5). Eines der in den Beispielen identifizierten Antigene evoziert sogar neutralisierende Antikörper und ist somit im therapeutischen Zusammenhang von großer Wichtigkeit. Somit ist das Verfahren auch für die Auffindung von für die Therapie wichtiger Antigene geeignet.

Das erfindungsgemäße Verfahren wurde in den vorliegenden Beispielen für die Identifikation von viralen und bakteriellen Antigenen eingesetzt. Erfindungsgemäß werden für bakterielle Erreger vorzugsweise die folgenden SDS-Konzentration zur Anreicherung von gramnegativen bzw. grampositiven Erregern eingesetzt (siehe Abb. 8): Für Gram-positive Bakterien >0,05% bis 1% und für Gram-negative Bakterien 0,05% bis 0,1% SDS.

Wenn keine Anhaltspunkte vorliegen, ob ein Erreger ein Virus oder ein Bakterium ist, kann aufgrund der sehr geringen benötigten Materialmenge die Ausgangsprobe aufgeteilt werden (z.B. auf zwei Probengefäße) und unter der Annahme eines ursächlichen viralen bzw. bakteriellen Erregers methodisch unterschiedlich verarbeitet werden (Abb. 12).

30

In einem weiteren Aspekt betrifft die vorliegende Erfindung neue Vaccinavirusantigene, dadurch gekennzeichnet, dass das Antigen von einer Nukleinsäure kodiert wird, die 80% Homologie, insbesondere 90% Homologie und vorzugsweise 95% Homologie zu einer der Sequenzen SEQ ID NOS: 4,5,6,7,8,9,10, 11,12, 13, 14, 15, 16 ,17, 18 ,19, 20, 21 oder 22 35 aufweist.

Besonders bevorzugt sind Vaccinavirusantigene, die dadurch gekennzeichnet sind, dass das Antigen von einer der Nukleinsäuren SEQ ID NOS: 4,5,6,7,8,9,10, 11,12, 13, 14, 15, 16 ,17, 18,19, 20, 21 oder 22 codiert wird.

Die erfindungsgemäßen Vaccinaantigene die durch das erfindungsgemäße Verfahren identifiziert wurden sind im Beispiel 5 und Tabelle 3 näher beschrieben. Bevorzugte Nukleinsäuresequenzen, die die Vaccinavirusantigene codieren sind im Sequenzprotokoll als SEQ ID NOS: : 4,5,6,7,8,9,10, 11,12, 13, 14, 15, 16 ,17, 18,19, 20, 21 oder 22 dargestellt. Die bevorzugten Nukleinsäuresequenzen sind zusätzlich in Abbildung 13 dargestellt..

5 In weiterer Aspekt der Erfindung betrifft daher auch die Verwendung der Nukleinsäuren SEQ ID NOS: 4-22 und Nukleinsäuren, die mit diesen 80, 90 oder 95% Homologie aufweisen in Nachweisverfahren des Vaccinavirus. Derartige Nachweisverfahren sind dem Fachmann bekannt.

10 Weitere Verwendungen für die erfindungsgemäßen Vaccinia virusantigene:

- Möglicher serologischer Nachweis von Variola major (Erreger der Pocken) über konservierte Epitope in Vaccinia virusantigenen. Variola major ist einer immunologischen Analyse nicht zugänglich, da in Militärlaboren verschlossen.
- Impfung gegen Variola major mit Subunitvakzinen. Die Impfung mit dem Vaccinia virus erzeugt einen guten Schutz gegen den Pockenerreger. Einzelne Vaccinia virusantigene sind daher potenzielle Kandidaten für die Induktion von immunologischem Schutz gegen den Pockenerreger.

20

Die folgenden Abbildungen dienen der Erläuterung der Erfindung:

**Abb.1.** Schematische Darstellung der aufeinanderfolgenden Analyseschritte einer Ausführungsform des erfindungsgemäßen Verfahrens.

25 **Abb 2A Amplifikation unterschiedlicher Ausgangsmengen von Erregernukleinsäuren.** Klenow-getaggte Vaccinia virus-DNA wurde wie im Beispiel 3A erläutert einmalig für 35 Zyklen amplifiziert (PCR1) bzw. 1 µl für weitere 35 Zyklen reamplifiziert (Re-PCR1). Als Ausgangsmengen für die Amplifikation dienten 1 ng (Spur 1), 40 pg (Spur 2), 8pg (Spur 3), 0,8pg (Spur 4), bzw 0,08pg (Spur 5) Klenow-Enzym getaggte Vaccinia virus-DNA. Spur 6 stellt die Negativkontrolle ohne Zugabe von Vaccinia virus-DNA dar.

30 **Abb. 2B.** Amplifikation von Vaccinia virus-DNA; die Amplifikation der Vaccinia virus-DNA wurde durch eine Klenow-Enzym-Reaktion eingeleitet (links), für die Adaptor oligonukleotide mit degeneriertem 3'-Ende für ein sequenzunabhängiges Priming verwendet wurden. Darauf folgt die eigentliche PCR-Amplifikation mit Oligonukleotiden, die der Adaptorsequenz entsprechen. Die beschriebenen PCR-Bedingungen führten zu Fragmenten unterschiedlicher 35 Länge (200-2500 bp) (rechts).

**Abb. 3.** Abb. 3 zeigt das Immunoscreening und die Identifizierung des 39kDa Antigen Klon 3 (288-939) und des ATI Antigen Klon 1 (511-111). Anfänglich im Screening identifizierte Klone wurden zunächst oligoklonal unter Zunahme von benachbarten nicht-reaktiven Phagenplaques isoliert und nach Bestätigung monoklonalisiert.

- 5 **Abb. 4A.** Abb. 4A zeigt Klon 1 (288-688), Klon 2 (288-788) und Klon 3 (288-938), die für überlappende Bereiche aus dem 39kDa-Protein von Vacciniaivirus kodieren. Die Klone sind unterschiedlich immunaktiv.

**Abb. 4B.** Abb. 4B zeigt drei Klone, die für überlappende Bereiche des A-type inclusion protein (ATI) von Vacciniaivirus kodieren und gleiche Immunreakтивität aufweisen.

- 10 **Abb. 4C.** Abb. 4C zeigt drei Klone, die für überlappende Bereiche des plaque size /host range protein (ps/hr) von Vacciniaivirus kodieren und unterschiedlich immunaktiv sind.

- 15 **Abb. 5.** Abb. 5 zeigt die Verteilung der mit dem erfindungsgemäßen Verfahren identifizierten Klone im Vacciniaivirusgenom. Die identifizierten Antigene sind über das ganze Vacciniagenom verteilt, was auf eine gute Repräsentanz der Vacciniaivirusgene in der über das erfindungsgemäße Verfahren hergestellten Bibliothek schließen lässt.

- 20 **Abb. 6.** Abb. 6 zeigt die molekulare Analyse der Repräsentanz von zehn arbiträr ausgesuchten Vacciniaivirusgenen in der durch das erfindungsgemäße Verfahren amplifizierten Vacciniaivirus DNA. Zehn Genabschnitte aus dem Genom des Vacciniaivirus wurden per PCR synthetisiert. Je 10 ng der 317 bis 549 bp langen Genabschnitte wurden durch Gelelektrophorese in Agarosegelen aufgetrennt, über das Southern Blotting-Verfahren auf eine Nylonmembran überführt. Für die Herstellung der <sup>32</sup>P-markierten Sonde wurden in 10-20 pg Vacciniaivirus-DNA eingesetzt. Abb. 6A zeigt die Hybridisierung mit PCR-25 Fragmenten aus einer einzigen Re-PCR. Abb. 6B zeigt die Verbesserung der Repräsentanz indem gepoolte Fragmente aus mehreren wie im Beispiel 2 beschrieben variierten Re-PCRs für die Hybridisierung eingesetzt wurden. Die Hybridisierung der gepoolten amplifizierten DNA mit den geblotteten zehn zufällig ausgewählten Abschnitten (sichtbar als schwache bis deutliche Schwärzung) des Vacciniaivirusgenoms zeigt, dass alle Abschnitte in der amplifizierten DNA enthalten sind. Herauszustellen ist, dass selbst das Gen mit dem schwächsten Hybridisierungssignal (Spur 2, 94kDA A-Type inclusion protein, ATI) im Immunoscreening der Bibliothek 30-mal als Antigen identifiziert wurde (s. Tabelle 3).

- 30 **Abb. 7A.** Abb. 7A zeigt die Bestimmung des Serumtiters gegen das durch das erfindungsgemäße Verfahren klonierte immundominante 39 kDa Antigen des Vacciniaivirus. Serum von C57BL/6 Mäusen wurde am Tag 21 nach Infektion mit dem Vacciniaivirus gewonnen und wie angegeben verdünnt. Zur Produktion des Antigens wurden E. coli

Bakterien mit einem das 39 kDa Antigen kodierenden Lambaphagen infiziert. Die Reaktivität der Serumverdünnungen gegen das so rekombinant exprimierte 39 kDa Antigen wurde auf Nitrocellulosemembranen getestet. Für das hier verwendete Serum ist der Antikörpertiter >1:16000.

- 5 **Abb. 7B.** In Abbildung 7B ist der Verlauf des Antikörpertiters gegen das 39 kDa Antigen nach Infektion mit  $2 \times 10^6$  pfu Vacciniavirus oder mit  $2 \times 10^5$  pfu des Lymphocytären Choriomeningitisvirus (Stamm WE) dargestellt. Nicht infizierte Tiere (naiv) zeigen keine Reaktivität gegen das 39 kDa Antigen. Die hohe Spezifität der Reaktion ist auch durch die nur minimale Kreuzreakтивität mit dem Serum aus mit dem Lymphocytären Choriomeningitisvirus infizierten Mäusen (Tag 14) gezeigt.
- 10

- 15 **Abb. 8.** Abb. 8 zeigt die unterschiedliche Empfindlichkeit von eukaryontischen Zellen und gramnegativen bzw. grampositiven Bakterien. Für das Experiment wurden vergleichbare Zellvolumina von gramnegativen Bakterien (oben), grampositiven Bakterien (Mitte) bzw. eukaryontischen Zellen mit den angegebenen Konzentrationen von SDS inkubiert. Nicht lysierte korpuskuläre Strukturen wurden durch Zentrifugation pelletiert. Während eukaryontische Zellen bereits bei geringsten Konzentrationen vollständig von SDS lysiert werden (kein sichtbares Zellpellet, mikroskopisch keine sichtbaren Zellen) sind Bakterien widerstandsfähiger und können, da ihre korpuskuläre Integrität erhalten blieb, durch 20 Zentrifugation angereichert werden.
- 20

- 25 **Abb. 9.** Abb. 9 zeigt die Identifikation und molkulare Charakterisierung von putativen Antigenen des humanpathogenen Bakteriums *Tropheryma whippelii*. Interleukin-10 und Interleukin-4 deaktivierte humane Makrophagen wurde mit *T. whippelii* Bakterien enthaltendem Gehirnmaterial eines an der Whippleschen Erkrankungen verstorbenen Patienten inkubiert. Bakterienspezifische Gene wurden durch differenzielle Lyse und anschliessende DNA-Aufarbeitung isoliert und Bibliotheken nach dem erfindungsgemässen Verfahren konstruiert. Das Immunoscreening wurde mit Seren von *T. whippelii* infizierten Patienten durchgeführt. Die bioinformatische Analyse, d.h. der Abgleich mit öffentlich zugänglichen Sequenzdatenbanken, zeigt, dass bisher nicht bekannt bakterielle Antigene 30 durch das erfindungsgemässe Verfahren identifiziert wurden.
- 30

- 35 **Abbildung 10:** Isolation von Bakterien direkt aus dem Milz-Gewebe eines Patienten mit Whipple-Disease. Bakterien wurde aus einer kryopräservierten Milzprobe eines Patienten mit Whipple-Disease wie im Beispiel 9 dargestellt isoliert und fluoreszensmikroskopisch analysiert. Das Bild zeigt die Überlagerung die Aufnahme im Phasenkontrast (Nachweis korpuskulärer Partikel, oben) und nach Überlagerung mit blauen Fluoreszenzsignal (Nachweis DNA).

**Abbildung 11A:** Anreicherung von Erregernukleinsäuren direkt aus einer Patientenprobe wie in Beispiel 9 beschrieben.

**Abbildung 11B:** Amplifikation von direkt aus Patientenproben isolierten Erregernukleinsäuren. Die aus der Milzprobe angereicherte Bakterien-DNA wurde wie im

5 Beispiel 10 beschrieben amplifiziert (Spur 2). Spur 1 stellt die Positivkontrolle mit einer anderen DNA-Probe dar. Auf Spur 3 ist die Negativkontrolle ohne Zusatz von DNA aufgetragen.

**Abbildung 12:** Schema eines möglichen Vorgehens für die Identifikation von Erregerantigenen für den Fall, dass nicht bekannt ist, ob der Erreger ein Virus oder ein

10 Bakterium darstellt. Die Ausgangsprobe wird aufgeteilt und mittels unterschiedlicher, die Identifikation von bakteriellen bzw. viralen Erregern ermöglichen Verfahren verarbeitet.

**Abbildung 13:** Abbildung 13 zeigt die Nukleinsäuresequenzen der in Tabelle 3 aufgelisteten identifizierten Vacciniaivirusantigene. Die Nukleinsäuresequenzen entsprechen den Sequenzen SEQ ID NO: 4 bis SEQ ID NO: 22 im Sequenzprotokoll.

15

Die vorliegende Erfindung wird weiterhin durch die folgenden Beispiele in nicht-limitierender Weise veranschaulicht.

### Beispiele

20 **Beispiel 1**

#### **Isolierung von Erregernukleinsäuren aus Virus-infizierten Zellen:**

BSC40 Zellen wurden mit  $2 \times 10^6$  pfu Vacciniaviren infiziert. Infizierte Zellen wurden bei 37°C im CO<sub>2</sub>-Inkubator 24 h inkubiert und geerntet. Geerntete Zellen wurden anschließend homogenisiert und in gepuffertem Medium aufgenommen. Zur Trennung von Viruspartikeln

25 von Wirtszellfragmenten wurde das in Medium aufgenommene Zellysat anschließend mit Ultraschall behandelt. Grobpartikuläre Strukturen wurden durch Zentrifugation 15min x 3000 U/min pelletiert. Nach Zentrifugation wurde der Überstand abgenommen und das Pellet verworfen. Um korpuskuläre Partikel im Überstand zu präzipitieren, wurde 2 ml gekühlter Überstand mit 6% PEG6000/0,6M NaCl präzipitiert (1h Inkubation auf Eis) und nachfolgend

30 das Präzipitat mittels Zentrifugation bei 10000 x g für 10 min pelletiert. Da die vorhergehenden Schritte sowohl eine Abtrennung als auch eine Lyse kontaminierender Wirtszellen erbracht haben sollten, wurde jetzt nach Verwerfen des Überstandes das Präzipitat in 300 µl DNase/RNase Puffer aufgenommen und 30 min bei 37°C mit RNase und DNase verdaut. Hierdurch wurden die nun extrazellulär vorliegenden Nukleinsäuren der vorher 35 desintegrierten Wirtszellen eliminiert. Die Virusnukleinsäuren werden durch die intakte Viruskapsid vor den Nuklease geschützt und nicht degradiert. Viruspartikel wurden durch

vortexen mit 1 Volumen GITC-Puffer aufgeschlossen, was auch zur Inaktivierung der zugegebenen Nukleasen führt. Freigesetzte Erreger-DNA wurde mit Phenol/Chloroform extrahiert und mit 1 Isovolumen Isopropanol gefällt. Die gefällten Nukleinsäuren wurden mit 80% Ethanol gewaschen und in 20 µl H<sub>2</sub>O dest. aufgenommen.

5

### Beispiel 2

#### Infektion mit Vacciniaivirus und Serumgewinnung

Rekombinantes Vacciniaivirus mit dem Glykoprotein des Vesikulären Stomatitis Virus (VaccG) (Mackett et al.(1985), *Science* 227, 433-435.) wurde auf BSC40 Zellen gezüchtet 10 und die Viruskonzentration im Plaque-Assay bestimmt. C57BL/6 Mäuse (Institut für Labortierkunde, Universität Zürich) wurden i.v. mit  $2 \times 10^6$  pfu VaccG infiziert. Den Mäusen wurden 200-300 µl Blut an den Tagen 8, 16 und 30 nach Infektion entnommen und Serum wurde durch Zentrifugation gewonnen und bei -20°C gelagert.

Nach erfolgter Immunisierung der Mäuse wurde die erfolgreiche Induktion von anti-15 Vacciniaantikörpern im Neutralisationsassay gegen VSV getestet (Ludewig et al.. 2000. Eur J Immunol. 30:185-196), um den besten Zeitpunkt der Serumabnahme für die geplanten Analysen zu ermitteln. Wie in Tabelle 1 dargestellt, hat ab Tag 16 der Anstieg sowohl des Gesamt-Immunglobulin, als auch der IgG-Klasse ihr Maximum erreicht, so dass dieses Serum der Abnahmetage 16 und 30 verwendet werden konnte.

20

**Tabelle 1: Titerverlauf von Antikörpern nach Immunisierung mit Vacciniaivirus**

Tag post inoculationem	Gesamt-Immunglobulin (in 40xlog2)	IgG (in 40xlog2)
8	9	6
16	12	11
30	12	12

25 **Beispiel 3A Globale Amplifikation geringster Mengen an Erreger-Nukleinsäuren**

Die globale Amplifikation von genomischen Nukleinsäuren des Erregers ist ein essentieller Schritt in dem erfindungsgemässen Verfahren. Hierbei ist die Hauptherausforderung die sehr geringe Menge genomicscher Keimnukleinsäuren, die (ohne Vorkultur) aus infizierten Geweben isoliert werden, umfassend (d.h. möglichst alle Abschnitte des Genoms

miteinschliessend) zu amplifizieren. Zudem muss die amplifizierte DNA für das nachfolgende Screening exprimierbar und klonierbar sein. Während PCR-amplifizierte cDNA-Expressionsbibliotheken vielfach beschrieben, hergestellt und benutzt werden (z.B. Edwards et al., 1991) und z.T. als Kitlösung angeboten werden (z.B. SMART-cDNA library construction kit, Clontech), ist die Herstellung umfassender genomischer Bibliotheken ausgehend von geringen Mengen (<10 ng) von Erreger-nukleinsäuren bisher nicht beschrieben.

Daher war es notwendig für das erfundungsgemässe Verfahren ein DNA-Amplifikationsmodul zu entwickeln, welche die Herstellung von umfassenden genomischen Expressionsbanken aus Subnanogrammengen Material erlaubt. Die Methode wurde für das Vaccinia-virusgenom etabliert und ist ohne Modifikationen für alle DNA-kodierten Erreger und mit geringen Modifikationen auf RNA kodierte Pathogene übertragbar. Die Amplifikation der Vaccinia-virus-DNA wurde durch eine Klenow-Enzym Reaktion eingeleitet; für die Adaptor-Oligonukleotide mit degeneriertem 3'-Ende für ein sequenzunabhängiges Priming nach dem Random-Prinzip verwendet wurden. Hierauf folgen sequentiell zwei PCR-Amplifikationen mit Oligonukleotiden, die der Adaptorsequenz entsprechen. In einer Reihe von unabhängig durchgeführten Experimenten sind die Bedingungen für eine besonders effiziente Amplifikation erarbeitet worden. Nach Optimierung der Methode wurden zur Bestimmung der Sensitivität verschiedene Mengen Vaccinia-Virus DNA (u.a. 25 ng, 1 ng, 200 pg, 20 pg, 2 pg) mit 2 pMol Adaptor-N(6) (GATGTAATAACGAA[P2][P3]TTGGACTCATATANNNNN) gemischt, 5 min bei 95°C denaturiert und anschließend auf Eis abgekühlt. N steht hierbei für den degenerierten Primeranteil. Nach Vorbereitung eines Reaktionsansatzes mit Klenow-Enzym (2 U), DNA-Polymerase-1 Puffer (10 mM Tris-HCl pH 7,5, 5 mM MgCl<sub>2</sub>, 7,5 mM Dithiothreitol, 1 nMol dNTPs) wurde eine Primerverlängerung bei 37°C für 2 h durchgeführt. Anschließend erfolgte die Reinigung der durch Klenow-Polymerase elongierten Fragmente von den freien Adaptoroligonukleotiden über Standardtechniken. Jeweils 1/25 (d.h. 1 ng, 40 pg, 8 pg, 0,8 pg, 0,08 pg) der mit Klenow-Polymerase getaggten DNA wurden für einen ersten Amplifikationsschritt eingesetzt. Die PCR-Amplifikation mit Adaptor-Oligonukleotiden wurde mit zwei unterschiedlichen Oligonukleotiden (EcoR1-Adaptoroligonukleotid GATGTAATAACGAATTGACTCATAT bzw. Mfe1-Adaptor-Oligonukleotid GATGTAATAACAATTGGACTCATAT) (Anlagerung bei 60°C 1 min; Verlängerung bei 72°C 2,5 min; Denaturierung 94°C 1min; 35 Zyklen) durchgeführt. Die einmalige Amplifikation der Nukleinsäuren bei Nukleinsäuren von weniger als 40 pg erwies sich als nicht ausreichend um eine optisch im Ethidiumbromid/Agarosegel detektierbaren Amplifikationsschmier zu erzeugen. Je 1 µl des Amplifikats wurden daher anschließend als Template in eine zweite Amplifikation unter identischen Bedingungen für 30-35 Zyklen überführt. Die amplifizierten Produkte wurden gelelektrophoretisch analysiert. Die beschriebenen PCR-Bedingungen führten zu einer Amplifikation mit Fragmenten unterschiedlicher Länge (150-2000 bp) in allen Versuchsanordnungen bis minimal 0,8 pg Template DNA (Abb. 2A). Dabei wurden bei niedrigeren Ausgangsmengen von DNA im Durchschnitt kürzere Fragmente amplifiziert. In verschiedenen Experimenten wurden die Bedingungen für die Re-PCR variiert. Hierbei erwies sich das eine Variation der

- Pufferbedingungen (z.B. Mg-Konzentration) und Variation der eingesetzten Enzyme z.B. Stoffel-Fragment der Taq-Polymerase unterschiedliche Amplifikationsmuster erzeugten (s. Abbildung 6). Für die Reamplifikation werden lediglich 1/50 der initialen PCR verwendet. Somit kann die Re-Amplifikation unter 50 verschiedenen Bedingungen durchgeführt werden.
- 5 Das eine Variation der Amplifikationsbedingungen eine besonders gute umfassende globale Amplifikation erlauben, wurde in der Repräsentanz-Analyse im reversen Southern-Blot nachgewiesen (s. Abb. 6). Zur Überprüfung der Identität amplifizierter Fragmente wurden diese über Standardverfahren in den Bluescript Klonierungsvektor (Stratagene) ligiert und 20 Klone wurden sequenziert. 20 von 20 Sequenzen stimmten mit Vaccinia-virus-Sequenzen
- 10 überein, sodass eine Amplifikation von Artefakt-Sequenzen (z.B. polymerisierte Primersequenzen) ausgeschlossen werden konnte.

### Beispiel 3B: Herstellung einer genomischen Bibliothek

Die Herstellung einer Vaccinia-Bibliothek erfolgte durch Amplifikation von 20 pg einer mit Klenow-Polymerase getaggten Vaccinia DNA analog den im Beispiel 3A durchgeführten Bedingungen. Für die PCR-Amplifikation mit Adaptor-Oligonukleotiden wurden in getrennten Ansätzen mit zwei unterschiedlichen Oligonukleotiden (EcoR1-Adaptoroligonukleotid GATGTAATACGAATTGACTCATAT bzw. MfeI-Adaptor-Oligonukleotid GATGTAATACAATTGGACTCATAT) jeweils 1/50 der aufgereinigten 15 Fragmente (Anlagerung bei 60°C 1 min; Verlängerung bei 72°C 2,5 min; Denaturierung 94°C 1 min; 35-40 Zyklen) verwendet. Je 1 µl des Amplifikats wurden anschließend als Template in eine zweite Amplifikation unter identischen Bedingungen für 30 Zyklen überführt. Die amplifizierten Produkte wurden gelelektrophoretisch analysiert. Die beschriebenen PCR-Bedingungen führten zu einer Amplifikation mit Fragmenten unterschiedlicher Länge (200-2500bp) (Abb. 2B). Alle Kontrollreaktionen, in denen kein DNA-Template eingesetzt wurde, blieben negativ. Anschließend wurden die amplifizierten Produkte aufgereinigt, mit EcoR1- bzw. MfeI-Restriktionsenzymen verdaut und in Lambda-ZAP-Express-Vektor (EcoR1-Fragment, Stratagene) ligiert. Die Kombination zweier unabhängiger Restriktionsenzyme erhöht die Diversifikation und die Wahrscheinlichkeit, dass immundominante Bereiche nicht 20 durch interne Restriktionsenzymeschnittstellen zerstört werden und damit dem Nachweis entgehen. Nach Ligation der Nukleinsäurefragmente in die Vektoren wurden diese nach Standardtechniken in Lambda-Phagen verpackt. Dies erfolgte mit kommerziell erwerblichen 25 Packaging Extrakten entsprechend den Herstellerangaben (z.B. Gigapack Gold III, Stratagene). Die so entstandenen Lambdaphagen-Bibliotheken (SE mit EcoRI-Adaptoren, SM 30 für MfeI-Adaptoren) wurden ohne weitere Amplifikation im Immunoscreening analysiert.

**Beispiel 4****Immunoscreening und Identifikation von Antigenen**

Das Immunoscreening wurde wie bereits in Sahin et al. (1995), *Proc Natl Acad Sci USA* 92: 11810-3; und Türeci et al. (1997), *Mol Med Today* 3, 342-349 beschrieben durchgeführt.

5 Bakterien des *E.coli* K12-abgeleiteten Stammes XL1 MRF wurden in der exponentiellen Wachstumsphase geerntet, auf eine  $OD_{600}=0,5$  eingestellt und mit den Lambda-phagen der beschriebenen Expressionsbank infiziert. Die Anzahl der Plaque-bildenden Einheiten (plaque-forming-units, pfu) wurde derart eingestellt, dass eine Subkonfluenz der Plaques vorlag (z.B. ~5000 pfu/ 145 mm Petrischale). Unter Zusatz von TOP-Agar und IPTG wurde der  
10 Infektionsansatz auf Agarplatten mit Tetrazyklin ausplattiert. Bei der Übernachtkultur bei 37°C bildeten sich auf dem Bakterienrasen Phagenplaques. Jeder einzelne Plaque repräsentiert einen Lambda-Phagen-Klon mit der in diesen Klon inserierten Nukleinsäure und enthält gleichzeitig das von der Nukleinsäure kodierte rekombinant exprimierte Protein.

Nitrozellulosemembranen (Schleicher & Schüll) wurden aufgelegt, um Abklatschpräparate 15 der rekombinanten Proteine herzustellen (Plaque Lift). Nach Waschschriften in TBS-Tween und Blocken unspezifischer Bindungsstellen in TBS + 10% Milchpulver erfolgte die Inkubation im Serum des infizierten Wirtes über Nacht. Es wurde zu diesem Zweck gepooltes Serum der Infektionstage 16 und 30 verwendet und 1:100 – 1:1.000 verdünnt. Nach weiteren Waschschriften wurden die Nitrozellulose-Membranen mit einem gegen Maus-IgG 20 gerichteten, sekundären AP-konjugierten Antikörper inkubiert. Bindungen von Serumantikörpern an in Phagenplaques rekombinant exprimierte Proteine konnten auf diese Weise durch eine Farbreaktion sichtbar gemacht werden. Klone, die als reaktiv mit Wirtsseren identifiziert wurden, konnten auf die Kulturplatte zurückverfolgt und von dort das entsprechende Phagenkonstrukt monoklonal isoliert werden. Nach erneuter Ausplattierung 25 wurden solche positiven Klone bestätigt. Durch *in vivo* Exzision wurde der Lambda-phagenklon zum Phagemid rezirkularisiert (Sahin et al. (1995), *Proc Natl Acad Sci USA* 92: 11810-3).

In beiden Banken (SE und SM) wurden wie oben geschildert insgesamt 150.000 Klone gescreent. Zu diesem Zweck wurde das gepoolte Serum der infizierten Tiere vom Tag 16 und 30 nach Infektion in einer Verdünnung von 1:500 verwendet. Primär identifizierte Klone wurden zunächst oligoklonal unter Zunahme von benachbarten nicht-reaktiven Phagenplaques isoliert und nach Bestätigung monoklonalisiert (Abb. 3). Es konnten 26 (SE-Bank) bzw. 41 (SM-Bank) Klone isoliert werden, die mit dem Serum der immunisierten Tiere reaktiv waren.

Alle identifizierten Klone wurden zusätzlich mit Präimmunseren von Mäusen desselben 35 Stammes getestet und waren nicht reaktiv.

**Tabelle 2: Anzahl reaktiver Klone nach Screening der SE- und SM-Bank**

<b>Bank</b>	<b>gescreente Klone</b>	<b>reaktive Klone</b>
Bank SE	150.000	26
Bank SM	150.000	41

- 5 Die Sequenzierung und der Datenbankvergleich deckte unter anderem die drei folgenden unterschiedlich immunogenen Vacciniaivirusproteine unter den Klonen auf:

#### 39 kDa immunodominant antigen protein

- 10 Drei Klone kodieren für Bereiche aus dem 39-kDa Protein von Vacciniaivirus (Abb.4A). Die Klone stellen Fragmente dieses Proteins dar. Sie beginnen alle bei Nukleotid 288, reichen aber unterschiedlich weit an das 3'-Ende dieses Genproduktes, nämlich bis Nukleotid 688, 788 bzw. 938.

15 Das für das 39-kDa Protein kodierende Gen ist der ORF A4L im Western Reserve (WR) Stamm (Maa und Esteban (1987), *J. Virol.* 61, 3910-3919). Das 281 Aminosäuren lange 39-kDa Protein ist in Menschen und Tieren stark immunogen (Demkovic et al. (1992), *J. Virol.* 66, 386-398).

20 Es wurde bereits beschrieben, dass eine Immunisierung mit dem 39-kDa Protein eine schützende Immunität in Mäusen induzieren kann (Demkovic et al. (1992), *J. Virol.* 66, 386-398).

25 Die stärkste antigene Domäne scheint innerhalb der letzten C-terminal gelegenen 103 Aminosäuren zu liegen (Demkovic et al. (1992), *J. Virol.* 66, 386-398). Die Lage der hier gefundenen Fragmente ist ebenfalls als Hinweis auf Sero-Epitope zu verstehen. Interessanterweise decken die beiden stark immunreaktiven Klone 2 und 3 den Bereich dieser als stark antigen beschriebenen 103 Aminosäuren ab.

30 An diesem Beispiel wird auch die Multidimensionalität der Aussagen des erfindungsgemäßen Verfahrens deutlich. Neben der Identifizierung des gleichzeitig auch Immunschutz gewährenden Antigens ist durch die Anzahl überlappender Klone ein Hinweis auf die Abundanz der Antikörper gegeben. Die Lage der Klone erlaubt eine Einengung der Sero-Epitope sowie die Stärke der Reaktivität einen Hinweis auf die Avidität der Antikörper. Dies gilt ebenso auch für die im folgenden beschriebenen Antigene.

### A-type inclusion Protein (ATI)

Einige der hier gefundenen Klone repräsentieren das A-type inclusion protein (ATI) (Abb. 4B), ein ca. 160 kDa großes Protein bei verschiedenen Orthopoxviren (Patel et al. (1986),

5 *Virology 149*, 174-189), das einen großen Anteil der Proteinmenge der charakteristischen Einschlusskörperchen ausmacht. Beim Vacciniaivirus ist dieses Protein trunkiert und nur ca. 94 kDa groß (Amegadzie (1992), *Virology 186*, 777-782). ATI assoziiert spezifisch mit infektiösem intrazellulären reifen Vacciniapartikeln und ist nicht in behüllten extrazellulären Vacciniaviren zu finden (Uleato et al. (1996), *J. Virol.* 70, 3372-3377). ATI ist eines der 10 immundominanten Antigenen in Mäusen, wobei die immundominanten Domänen am Carboxyterminus des Moleküls liegen (Amegadzie et al. (1992), *Virology 186*, 777-782). Die drei hier gefundenen Klone, die identische Reaktionsstärken haben, decken den Bereich zwischen bp 308 und 1437 ab und sind entsprechend tatsächlich C-terminal bis zentral im kodierten Protein gelegen.

15

### Plaque size/host range (ps/hr) Protein

Das 38 oder 45 kDa große plaque size/host range Protein (ps/hr) wird durch den ORF B5R kodiert (Takahashi-Nishimaki et al. (1991), *Virology 181*, 158-164). ps/hr ist ein Typ 1

20 Transmembranprotein, das in die Membran von extrazellulären Viruspartikeln inkorporiert wird oder von Zellen während der Infektion sezerniert werden kann. Antikörper gegen ps/hr neutralisieren die Infektiösität des Vacciniaivirus (Galmiche et al. (1999), *Virology 254*, 71-80). Deletion von ps/hr führt zur Attenuierung des Virus *in vivo* (Stern et al. (1997), *Virology 233*, 118-129). Ausserdem schützt die Immunisierung mit B5R gegen eine Infektion mit ansonsten tödlichen Dosen des Virus (Galmiche et al. (1999), *Virology 254*, 71-80). Drei der 25 hier identifizierten Klone stellen Fragmente dar, die wiederum denselben Bereich dieses Antigens abdecken und den C-Terminus einschließen (Abb. 4C). Dies bedeutet, dass das von diesen Klonen repräsentierte Sero-Epitop im extrazellulären Bereich dieses viralen Oberflächenmoleküls liegt und somit gut für Antikörper zugänglich ist.

30 **Beispiel 5**

### Sequenzierung und bioinformatische Analyse der identifizierten Vacciniaivirusantigene

Die Sequenzierung der identifizierten Klone erfolgte nach Standardtechniken mit Oligonukleotiden, die das Insert flankieren (BK-Reverse, BK-Universe) in Sanger'schen Kettenabbruchverfahren. Ermittelte Sequenzen wurden über BLAST-Analyse mit bekannten

35 Sequenzen in der Genbank abgeglichen. Die Lokalisation der Vacciniaivirusantigene im Genom (Accessionsnummer M35027) und die Standardnomenklatur ist in Tabelle 3 angegeben. Diese Analyse zeigt, dass über das gesamte Vacciniaivirusgenom verteilte

Antigene mit dem erfindungsgemässen Verfahren identifiziert wurden. Bei einer grossen Anzahl der identifizierten Gene ist bisher nicht bekannt gewesen, dass die Genprodukte eine Wirkung als Antigen besitzen. Mit Hilfe des erfindungsgemässen Verfahrens werden daher nicht nur bekannte Antigene, sondern auch unbekannte Antigene 5 identifiziert. Ein weiterer Vorteil des erfindungsgemässen Verfahrens ist, dass Antigene identifiziert werden können, die auf beiden Strängen (kodierender und komplementärer Strang) des Genoms zu finden sind.

10

**Tabelle 3: Identität, genomische Lokalisation, serologische Reaktivität und Anzahl der identifizierten Vacciniaivirusantigene mit Hilfe des erfindungsgemässen Verfahrens.**

Vacciniaivirus Antigene	Lokalisation in VV-Genom	SEQ ID NO:	Signal	# Klonen
39 kDa immunodominant antigenic antigen (A4L)	ORF 151 (117270-116425)	4	++++	62
94 kDa A-type inclusion protein (TA31L)	ORF 174 (138014-135837)	5	+++	30
35 kDa plaque size/host range protein (B5R)	ORF 232 (167383-168336)	6	+++	7
116 kDa DNA polymerase (E9L)	ORF 80 (59787-56767)	7	+++	4
65 kDa envelope protein (F12L)	ORF 60 (43919-42012)	8	+++	2
62 kDa rifampicin resistance gene (D13L)	ORF 145 (113026-111371)	9	+++	1
32 kDa carbonic anhydrase-like protein (D8L)	ORF 137 (107120-106206)	10	+++	1
36 kDa late protein (I1L)	ORF 87 (63935-62997)	11	+++	1
16 kDa protein (TC14L)	ORF10 (10995-10567)	12	+++	1
38 kDa serine protease inhibitor 2 (B13R)	ORF 421 (172562-172912)	13	++	4
18 kDa protein (C7L)	ORF 24 (19257-18805)	14	++	3
24.6 kDa protein (B2R)	ORF 226 (163876-164535)	15	++	2
36 kDa protein (A11R)	ORF 164 (124976-125932)	16	++	1
15 kDa membrane phosphoprotein (A14L)	ORF 167 (126785-127128)	17	++	1
147 kDa protein (J6R)	ORF 117 (86510-90370)	18	++	1
77 kDa protein (O1L))	ORF 84 (62477-60477)	19	++	1
59 kDa protein (C2L)	ORF30 (24156-22618)	20	+	4

90 kDa protein (DSR)	ORF132 (101420-103777)	21	+	1
23 kDa protein (A17L)	ORF170 (129314-128703)	22	+	1

In der Abbildung 5 ist eine grafische Repräsentation des Vacciniaivirusgenoms mit Darstellung der offenen Leseraster (ORF) zu finden, die zeigt, dass die durch das erfindungsgemäße Verfahren identifizierten Antigene über das gesamte Vacciniaivirusgenom verteilt sind. Dies indiziert, dass das erfindungsgemäße Verfahren eine repräsentative Amplifikation von spezifischer Erregernukleinsäure aus geringsten Mengen von Ausgangsmaterial (1 – 20 pg) erlaubt.

### Beispiel 6

10 **Repräsentationsanalyse durch reversen Southern Blot**

Die repräsentative Amplifikation aus geringsten Mengen von Erregernukleinsäuren durch das erfindungsgemäße Verfahren wurde ebenfalls im folgenden Experiment gezeigt.

Zehn Genabschnitte des Vacciniaivirusgenoms wurden ausgewählt und durch PCR-Reaktionen amplifiziert. Nach gelelektrophoretischer Auf trennung wurden die DNA-Fragmente über alkalischen Transfer auf Nylonmembranen geblottet. Radioaktive Hybridisierung wurde mit 20 ng nach dem erfindungsgemäßen Verfahren hergestellter und mit <sup>32</sup>P markierter DNA durchgeführt. Abb. 6A zeigt, dass nur ein Teil der zehn zufällig ausgewählten Abschnitte des Genoms in der nach dem erfindungsgemäßen Verfahren hergestellten DNA in einer einzelnen Re-PCR DNA enthalten sind. Werden mehrere Re-PCR DNA, die in unterschiedlichen Ansätzen und unter varierten Bedingungen hergestellt wurden, kombiniert, ist die 100%ige Repräsentanz der zufällig ausgewählten Genabschnitte in der nach dem erfindungsgemäßen Verfahren hergestellten DNA offensichtlich (Abb. 6B). Die unterschiedliche Abundanz der Nukleinsäuren, z.B. des 39 kDa Antigens (Abb. 6B, Spur 10), kann zumindest teilweise erklären, dass bestimmte Genabschnitte häufiger in der nach dem erfindungsgemäßen Verfahren hergestellten DNA gefunden werden. Eine niedrige Abundanz der DNA im Amplifikat schliesst aber eine häufige Detektion im Screening nicht aus, wie das Beispiel der A-type inclusion protein DNA (Abb. 6B Spur 2) zeigt.

### Beispiel 7

#### Differentielle Serologie

30 Lambdaphagen, deren rekombinante Inserts für Antigene kodierten, die durch Antikörper im Serum infizierter Mäuse erkannt werden, wurden auf Reaktivität mit Seren nicht infizierter Tiere (immunologisch naiv) und mit Seren von mit dem Lymphocytären Choriomenigitisvirus infizierten Mäusen getestet. Diese Studien wurden ebenfalls als Plaque-Lift Assay durchgeführt. Der Antikörpertiter und die spezifische Reaktivität der Seren gegen

Lift Assay durchgeführt. Der Antikörpertiter und die spezifische Reaktivität der Seren gegen die klonierten Antigene kann so einfach ermittelt werden. In Abb. 7A ist dargestellt, wie die Reaktivität eines am Tag 21 nach Infektion mit dem Vaccinia-virus gewonnenen Serum gegen das 39 kDa Antigen ermittelt wurde. Zweifache Serumverdünnungen wurden mit durch Phagen in *E. coli* induziertes, rekombinantes 39 kDa Antigen inkubiert. Eine spezifische Reaktivität ist bei einer Serumverdünnung von 1:16000 noch erkennbar. Der zeitliche Verlauf der Antikörperreakтивität gegen das 39 kDa Antigen in mit Vaccinia-virus infizierten, Lymphocytären Choriomeningitisvirus infizierten und in nicht infizierten Mäusen ist in Abb. 7B dargestellt. Der Verlauf der Antikörperantwort nach Vaccinia-virusinfektion ist für diese Infektion typisch. Das Fehlen einer Reaktivität gegen das 39 kDa Antigen in naiven Mäusen und die nur sehr geringe Kreuzreaktivität nach Infektion mit dem Lymphocytären Choriomeningitisvirus zeigt die gute diagnostische Qualität, die mit durch das erfundungsgemäße Verfahren identifizierten Antigenen erreicht werden kann.

### Beispiel 8

#### 15 Identifikation von bakteriellen Antigenen mit Hilfe des erfundungsgemäßen Verfahrens

Zunächst wurden die Bedingungen erarbeitet, die eine Anreicherung von bakteriellen Erregern aus infizierten Proben erlauben. Hierzu wurde der Umstand genutzt, dass Bakterienwände widerstandsfähig gegenüber einer Lyse mit Solventien wie z.B. SDS sind. Durch eine SDS-Konzentrationsreihe wurde die Stabilität vom gramnegativen, grampositiven 20 Bakterien und eukaryontischen Zellen bestimmt. Wie in Abbildung 8 dargestellt bleibt unter 1% SDS die korpuskuläre Struktur grampositiver Bakterien erhalten. Gramnegative Bakterien wie in diesem Beispiel *E. Coli* weisen immerhin eine Widerstandsfähigkeit gegen Lyse bis auf 0,1% SDS auf. Dagegen werden alle Membranstrukturen (Zytoplasma, Kern) von eukaryontischen Zellen, wie in diesem Fall Fibroblasten, vollständig aufgelöst. Die hohe 25 SDS-Sensitivität von eukaryontischen Zellen wurden in anderen Beispielen auch für Leukozyten, Milzzellen, Lymphknotenbiopsien verifiziert. Nach Erarbeitung der Bedingungen wurde das erfundungsgemäße Verfahren für einen bisher nur unzureichend charakterisierten Erreger, *Tropheryma whippelii*, durchgeführt. *Tropheryma whippelii* ist ein 30 gram-positives Bakterium und die Infektion mit diesem Erreger kann die Whipplesche Erkrankung auslösen. Die Whipplesche Erkrankung ist eine chronische Infektion verschiedener Organe mit der Hauptmanifestation im Darm, die unerkannt zum Tode führen kann. Dieser Erreger ist nur schwer *in vitro* anzüchtbar, so dass für molekulare Analysen nur geringste Mengen an spezifischer Nukleinsäure zur Verfügung stehen. Aufgrund dieser Probleme war die Analyse der antigenen Strukturen dieses Erregers bisher nicht möglich.

35 Durch das erfundungsgemäße Verfahren gelang es, potentielle Antigene dieses Erregers zu definieren. In der Abb 9 sind die wesentlichen Schritte, die zur Charakterisierung von *Tropheryma whippelii* spezifischen Antigenen geführt haben, dargestellt.

*T. whippelii* haltiges, homogenisiertes Gehirnmaterial eines an der Whippleschen Erkrankung verstorbenen Patienten wurde verwendet, um mit Interleukin-10 und Interleukin-4

deaktivierte Makrophagen zu inokulieren (Schoedon et al. (1997) J Infect Dis. 176:672-677). Am Tag 7 postokulationem wurden infizierte Makrophagen geerntet, und 25 µl des Makrophagen/Bakterien Gemischs wurden verarbeitet. Die differenzielle Zellyse erfolgte durch Inkubation der mit Bakterien infizierten Makrophagen in 1% SDS enthaltendem Proteinase K Puffer für 15 min mit 20 µg/ml Proteinase K bei 55°C. Durch diese Behandlung wurden die in der Mischung vorhandenen Makrophagen (eukaryontische Zellen) aufgelöst. Durch die Lyse der Makrophagen werden deren Nukleinsäuren (RNA, DNA) in die Lösung freigesetzt. Dagegen verbleiben die Nukleinsäuren der Bakterien durch Erhaltung der Integrität der grampositiven Bakterienwand in den Bakterienzellen. Nach der Inkubation mit SDS/Proteinase K wurden die Bakterien durch Zentrifugation der Suspension pelletiert. Dagegen liessen sich die Bakterien bei fehlendem Zusatz von Proteinase K aufgrund hoher Viskosität der Lösung schlechter pelletieren. Der Überstand mit den Nukleinsäuren der Makrophagen wurde verworfen und das kaum sichtbare Pellet mehrfach gewaschen. Nach den Waschschritten wurden pelletierte Bakterien in 100 µl Wasser resuspendiert und jeweils 10 µl der Suspension lichtmikroskopisch und nach Färbung mit dem DNA-Farbstoff DAPI in der Immunfluoreszenz zur Bestimmung der Bakterienzahl verwendet. Die Anzahl der aus 25 µl der infizierten Makrophagen isolierten Bakterien betrug nach mikroskopischer Auszählung ca. 4000-6000 DNA-haltige Partikel. Die residualen 80 µl angereicherten Bakterien wurden zur Gewinnung von bakteriellen Nukleinsäuren nach Standardtechniken (Kochen, Denaturierung, DNA-Isolation mit Phenol/Chloform) weiterverarbeitet. Die Menge der aus den Bakterien isolierten DNA konnte aufgrund der niedrigen Menge nicht experimentell quantifiziert werden (kein detektierbares Signal im EtBr-Gel). Aufgrund der licht- und immunfluoreszenzmikroskopisch ermittelten Bakterienzahl (max. 6000) wurde bei einer angenommenen Bakteriengenomgrösse von 1-2 Mio Basen doppelsträngige DNA eine maximalen Ausbeute von 6-12 pg Erreger-DNA kalkuliert (Für die Kalkulation wurde das Durchschnittsgewicht eines Nukleotids von 660 eingesetzt). 50% der extrahierten DNA (d.h. max. 3-6 pg) wurde wie nach dem erfindungsgemässen Verfahren beschrieben (Klenow, sequentielle PCR, Re-PCR) amplifiziert und aus den amplifizierten Fragmenten eine genomische Bibliothek in Lambda-ZAP-Express-Vektor (Stratagene) hergestellt. Das Immunoscreening wurde mit Seren von *T. whippelii* infizierten Patienten durchgeführt. Positive Klone wurden sequenziert und bioinformatisch analysiert. Als Beispiel ist in Abb. 8 ein Klon aufgeführt, der sowohl für ein bakterielles putatives Lipoprotein als auch für ein putatives Histidine triad protein kodiert.

Dieses Beispiel zeigt, dass das erfindungsgemässen Verfahren neben der Identifikation von viralen Antigenen auch geeignet ist, bakterielle Antigene zu identifizieren.

#### **Beispiel 9: Anreicherung von Whipple Bakterien aus einer infizierten Milzprobe eines Patienten mit systemischer Infektion.**

Jeweils 20 µl einer kryopräservierten Milzprobe eines Patienten mit Morbus Whipple wurden unter fünf leicht modifizierten Bedingungen zur Anreicherung von Bakterien verwendet

(insgesamt 100 µl). Hierzu wurden die Milzproben wie im Beispiel 8 beschrieben in 1,5 ml Proteinase K Puffer mit 20 mg/ml Proteinase K Zusatz 10-60 min bei 55°C inkubiert. Danach wurden die in den infizierten Milzprobe befindlichen Bakterien wie im obigen Beispiel beschrieben durch Zentrifugation angereichert und mikroskopisch wie vorher beschrieben dokumentiert. Abbildung 10 zeigt eine Aufnahme einer bakterienreichen Pelletfraktion. Anschliessend wurden die Bakterien durch Kochen in GITC-Puffer und Ultraschallbehandlung aufgeschlossen und die bakteriellen Nukleinsäuren über Standardverfahren isoliert. Wie erwartet, war die Menge der Nukleinsäuren, die aus den angereicherten Fraktionen isoliert wurden, unter der Nachweisgrenze von 1 ng. Zur Dokumentation der bakteriellen Anreicherung wurden jeweils 1/100 der isolierten Nukleinsäuren für eine PCR-Ampifikation mit Whipple-Bakterien (Sequenz noch einfügen) bzw. human-DNA (Sequenz noch einfügen) spezifischen Oligonukleotiden eingesetzt und über 37 Zyklen amplifiziert. Abbildung 11A zeigt die Amplifikations-Resultate (A: PCR spezifisch für Whipple Bakterien, B: PCR spezifisch für humane DNA). Die Ergebnisse sind dargestellt für Amplifikationsbanden der Bakterien-angereicherten Fraktionen (Spur 1-3) bzw. von aus der nicht-angereicherten Fraktionen (Spur 4-6). Während in den nicht-angereicherten Fraktionen wie erwartet die Amplifikationssignale für humane DNA deutlich stärker sind (Spur 4-6) zeigen besonders Fraktionen 1 und 2 fast ausschliesslich eine Amplifikation von Erreger-nukleinsäuren. Das Beispiel zeigt das die geringe Menge des benötigten Materials es ermöglicht, die Anreicherungsbedingungen leicht zu variieren und anschliessend mit der am stärksten angereicherten Erreger-Fraktion (in diesem Fall 1 und 2) das erfindungsgemässe Verfahren fortzusetzen.

**Beispiel 10:** Globale Amplifikation von Erreger-nukleinsäuren aus direkt aus einer Milzprobe isolierten DNA. Die nach dem erfindungsgemäßen Verfahren wie im Beispiel 9 dargestellt isolierten Erreger-nukleinsäuren wurden wie im Beispiel 3 beschrieben amplifiziert und aus den amplifizierten Fragmenten eine genomische Bibliothek in Lambda-ZAP-Express-Vektor (Stratagene) hergestellt. Das Immunoscreening wurde mit Seren von *T. whippelii* infizierten Patienten durchgeführt. Positive Klone wurden sequenziert und bioinformatisch analysiert.

PATENTANSPRÜCHE

1. Verfahren zur Identifizierung von durch das Genom von mikrobiellen Erregern kodierten biologisch aktiven Strukturen ausgehend von genomischen Erreger-nukleinsäuren, umfassend die Schritte
  - (a) Gewinnung von genomischen Erreger-nukleinsäuren aus erregerhaltigen Proben,
  - (b) Sequenzunabhängige Amplifikation von genomischen Erreger-nukleinsäuren,
  - (c) Expression amplifizierter Erreger-nukleinsäuren, und
  - (d) Screening und Identifizierung der biologisch aktiven Struktur.
2. Verfahren gemäß Anspruch 1, dadurch gekennzeichnet, dass die durch das Genom von mikrobiellen Erregern kodierten biologisch aktiven Strukturen durch das Genom eines **bakteriellen Erregers kodiert werden**.
3. Verfahren gemäß Anspruch 1, dadurch gekennzeichnet, dass die durch das Genom von mikrobiellen Erregern kodierten biologisch aktiven Strukturen durch das Genom eines **DNA-haltigen Virus kodiert werden**.
4. Ein Verfahren gemäß Anspruch 1-3, wobei der mikrobielle Erreger ein intrazellulärer viraler oder bakterieller Erreger ist.
5. Ein Verfahren gemäß Anspruch 1-3, wobei der mikrobielle Erreger ein extrazellulärer viraler oder bakterieller Erreger ist.
6. Ein Verfahren gemäß den Ansprüchen 1-3, wobei der mikrobielle Erreger nicht-vital und/oder nicht-infektiös ist.
7. Ein Verfahren gemäß Anspruch 1, wobei die biologisch aktive Struktur ein Erregerantigen ist.

8. Ein Verfahren gemäß Anspruch 1, wobei die biologisch aktive Struktur ein Pathogenitätsfaktor des mikrobiellen Erregers ist.

9. Ein Verfahren gemäß Anspruch 1, wobei die biologisch aktive Struktur ein enzymatisch aktives Protein ist.

10. Ein Verfahren gemäss Anspruch 1, wobei zur Identifikation 10-20 pg Erregernukleinsäure verwendet werden.

10 11. Ein Verfahren gemäss Anspruch 1, wobei zur Identifikation 1-10 pg Erregernukleinsäure verwendet werden.

12. Ein Verfahren gemäss Anspruch 1, wobei die Proben Blut, Gewebe, kultivierte Zellen, Serum, Sekret aus Läsionen und andere Körperflüssigkeiten umfassen.

15

13. Ein Verfahren gemäss Anspruch 1, wobei die Gewinnung von genomischen Erregernukleinsäuren aus erregerhaltigen Proben in Schritt (a) die folgenden Schritte umfasst:

(a<sub>1</sub>) Freisetzung von Erregerpartikeln aus erregerhaltigen Proben

optional (a<sub>2</sub>) Elimination und/oder Reduktion kontaminierender

20 Wirtsnukleinsäuren

und (a<sub>3</sub>) Extraktion der genomischen Erregernukleinsäure aus freigesetzten Erregerpartikeln.

25 14. Ein Verfahren gemäss Anspruch 13, wobei die Freisetzung von Erregerpartikeln in Schritt (a<sub>1</sub>) durch Zellyse, Sedimentation, Zentrifugation und/oder Filtration erfolgt.

30 15. Ein Verfahren gemäss Anspruch 13, wobei die Elimination und/oder Reduktion kontaminierender Wirtsnukleinsäuren in Schritt (a<sub>2</sub>) durch RNase- und/oder DNase-Verdau erfolgt.

16. Ein Verfahren gemäss Anspruch 13, wobei die Extraktion der genomischen Erregernukleinsäure in Schritt (a<sub>3</sub>) durch Abtrennung des genetischen Materials des Erregers

von korpuskulären Bestandteilen des Erregers durch Proteinase K-Verdau, Denaturierung, Lysozym-Behandlung oder organische Extraktion erfolgt.

17. Ein Verfahren gemäss Anspruch 1, wobei die sequenzunabhängige  
5 Amplifikation der genomischen Erregernukleinsäure in Schritt (b) durch Klenowreaktion mit Adaptoroligonukleotiden mit degeneriertem 3'-Ende und nachfolgender PCR mit Oligonukleotiden, die der Adaptorsequenz entsprechen, erfolgt.

18. Ein Verfahren gemäss Anspruch 1, wobei die Amplifikation der  
10 Erregernukleinsäure in Schritt (b) durch reverse Transkription mit degenerierten Oligonukleotiden und nachfolgender Amplifikation durch PCR erfolgt.

19. Ein Verfahren gemäss Anspruch 1, wobei die Amplifikation der Erregernukleinsäure in Schritt (b) durch reverse Transkription mit degenerierten  
15 Oligonukleotiden und nachfolgende Amplifikation mit T7 RNA Polymerase erfolgt.

20. Ein Verfahren gemäss Anspruch 1, wobei zur Expression von amplifizierten Erregernukleinsäuren in Schritt (c) ein Einbringen der Erregernukleinsäuren in Vektoren und Verpackung der Vektoren in Lambda-Phagen erfolgt.  
20

21. Ein Verfahren gemäss Anspruch 1, wobei zur Expression von amplifizierten Erregernukleinsäuren in Schritt (c) ein Einbringen der Erregernukleinsäuren in filamentöse Phagen-Vektoren erfolgt.

25 22. Ein Verfahren gemäss den Ansprüchen 20 und 21, wobei die Vektoren ausgewählt sind aus der Gruppe der viralen, eukaryotischen oder prokaryotischen Vektoren.

23. Ein Verfahren gemäss Anspruch 1, wobei das Screening ein Immunoscreening auf Erregerantigene darstellt und die Identifizierung von Erregerantigenen in Schritt (d) die  
30 folgenden Schritte umfasst:

- (d<sub>1</sub>) Infektion von Bakterien mit Lambda-Phagen,
- (d<sub>2</sub>) Kultivierung der infizierten Bakterien unter Bildung von Phagenplaques,

(d<sub>3</sub>) Transfer der Phagenplaques auf eine Nitrozellulosemembran oder einer anderen Festphase, die zur Immobilisierung von rekombinanten aus den Erregern abgeleiteten Proteinen geeignet ist,

5 (d<sub>4</sub>) Inkubation der Membran mit Serum oder antikörperhaltigen Körperflüssigkeiten des infizierten Wirtes,

(d<sub>5</sub>) Waschen der Membran,

(d<sub>6</sub>) Inkubation der Membran mit sekundärem AP-gekoppeltem anti-IgG-Antikörper der spezifisch für Immunglobuline des infizierten Wirtes ist.

(d<sub>7</sub>) Nachweis der mit Wirtsserum reaktiven Klone durch Farbreaktion, und

10 (d<sub>8</sub>) Isolierung und Sequenzierung der reaktiven Klone.

24. Ein Verfahren gemäss Anspruch 1, wobei das Screening ein Immunoscreening auf Erregerantigene darstellt und die Identifizierung von Erregerantigenen in Schritt (d) die folgenden Schritte umfasst:

15 (d<sub>1</sub>) Generierung von rekombinanten filamentösen Phagen durch Einbringen der filamentösen Phagenvektoren in Bakterien,

(d<sub>2</sub>) Inkubation generierter rekombinanter filamentöser Phagen mit Serum eines infizierten Wirtes,

20 (d<sub>3</sub>) Selektion der filamentösen Phagen, an die Immunglobuline des Wirts gebunden haben, über immobilisierte Reagentien die spezifisch für die Immunglobuline des infizierten Wirts sind, und

(d<sub>4</sub>) Isolierung und Sequenzierung der selektierten Klone.

25 25. Ein Verfahren gemäss Anspruch 1, wobei der mikrobielle Erreger vor Gewinnung der Nukleinsäuren durch Fällung mit Polyethylenglykol, Ultrazentrifugation, Gradientenzentrifugation oder durch Affinitätschromatographie angereichert wird.

26. Vaccinavirusantigen, dadurch gekennzeichnet, dass das Antigen von einer Nukleinsäure codiert wird, die 80% Homologie zu einer der Sequenzen SEQ ID NOS: 4,5,6,7,8,9,10, 11,12, 13, 14, 15, 16 ,17, 18 ,19, 20, 21 oder 22 aufweist.

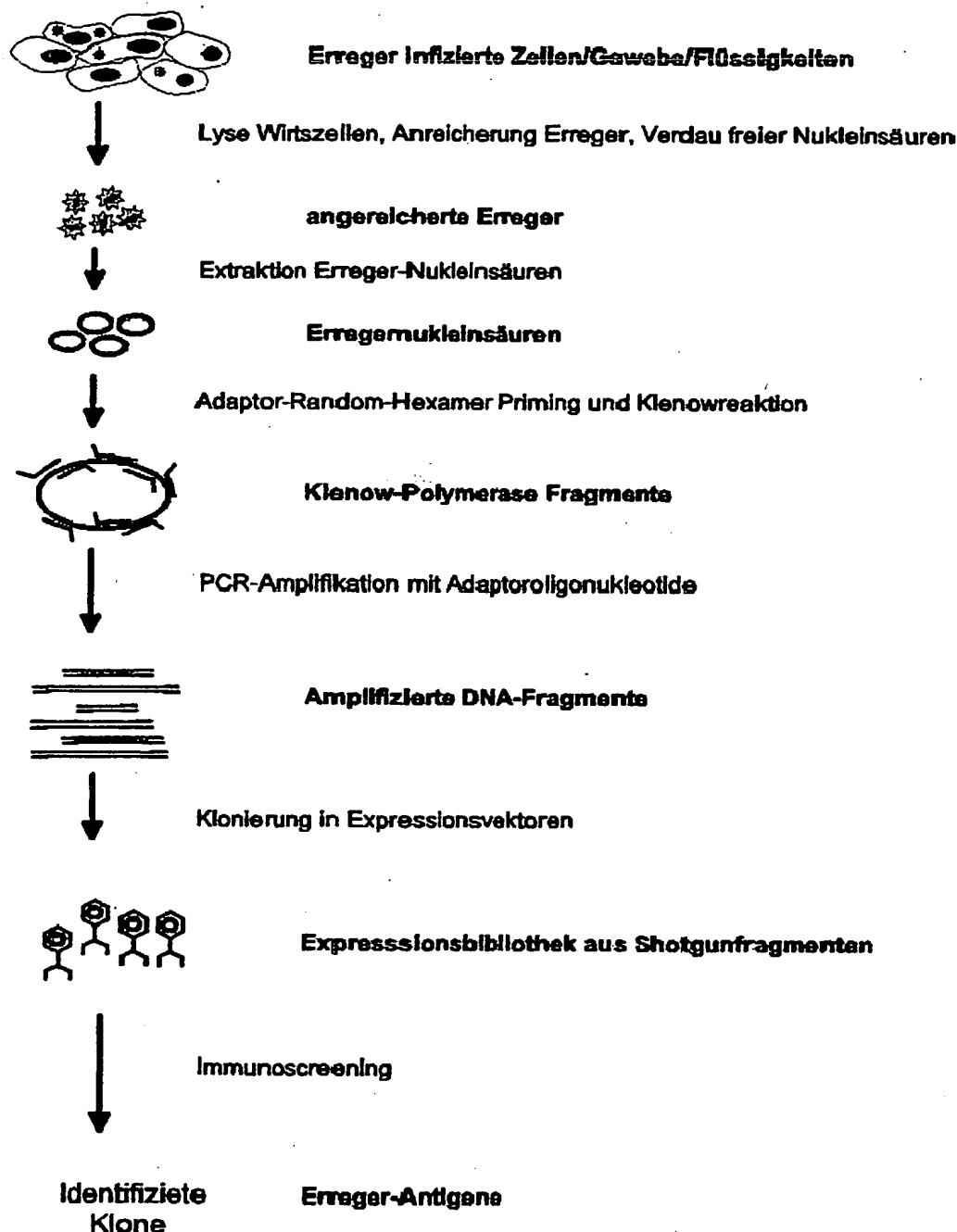
30 27. Vaccinavirusantigen nach Anspruch 26, dadurch gekennzeichnet, dass das Antigen von einer Nukleinsäure codiert wird, die 90% Homologie zu einer der Sequenzen SEQ ID NOS: 4,5,6,7,8,9,10, 11,12, 13, 14, 15, 16 ,17, 18 ,19, 20, 21 oder 22 aufweist.

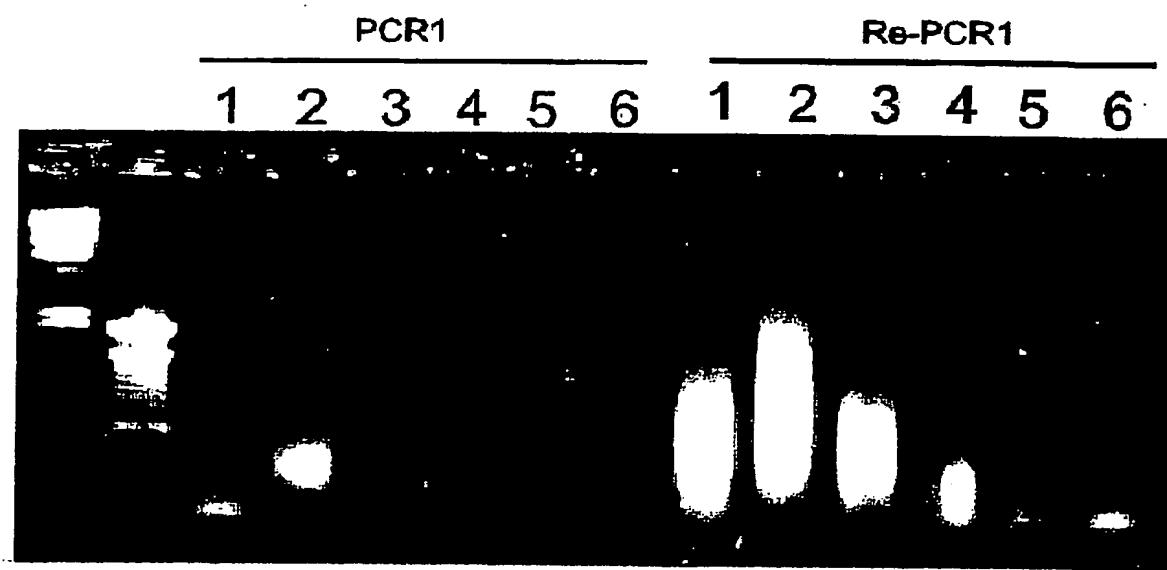
35 28. Vaccinavirusantigen nach Anspruch 26, dadurch gekennzeichnet, dass das Antigen von einer Nukleinsäure codiert wird, die 95% Homologie zu einer der Sequenzen SEQ ID NOS: 4,5,6,7,8,9,10, 11,12, 13, 14, 15, 16 ,17, 18 ,19, 20, 21 oder 22 aufweist.

34

29. Vaccinavirusantigen nach Anspruch 26, dadurch gekennzeichnet, dass das Antigen von einer der Nukleinsäuren SEQ ID NOS: 4,5,6,7,8,9,10, 11,12, 13, 14, 15, 16 ,17, 18 ,19, 20, 21 oder 22 codiert wird.

Fig. 1





**Abbildung 2A**

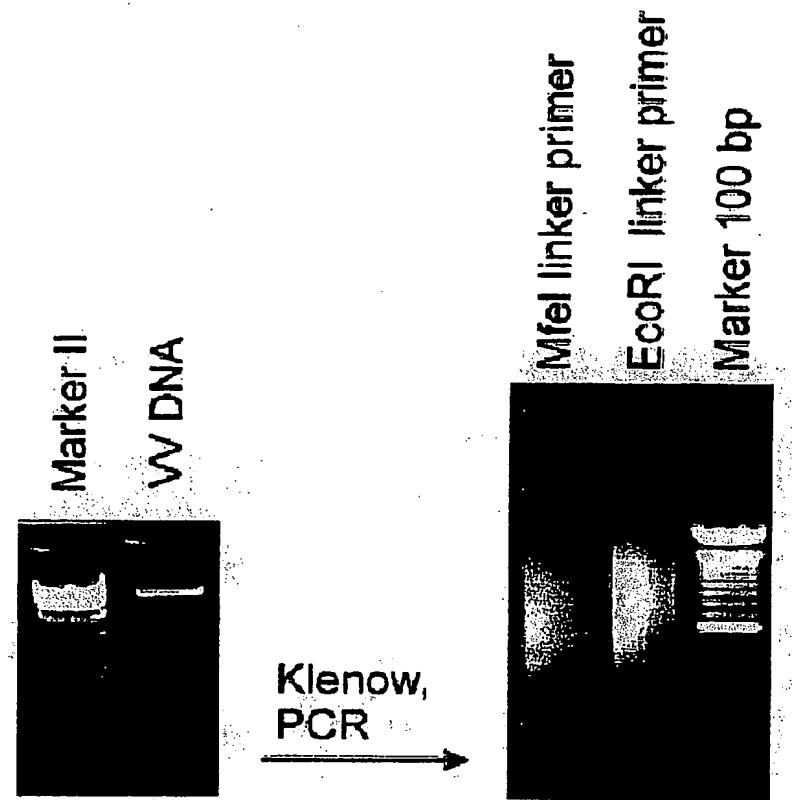
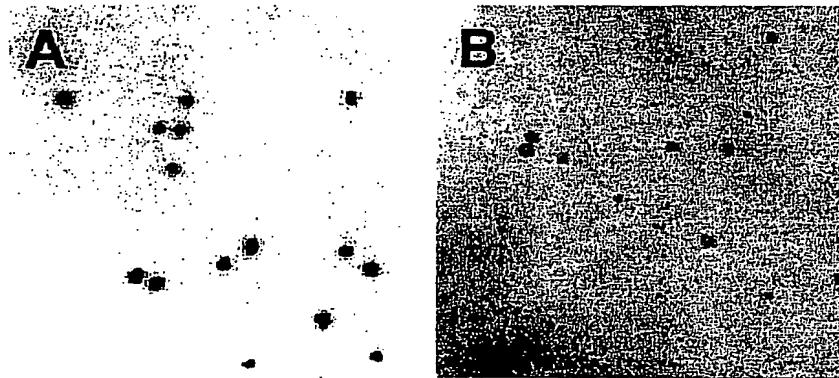


Abbildung 2 B



**39 kDa antigen  
Clone 3 (288-938)**

**ATI antigen  
Clone 1 (511-1111)**

**Abbildung 3**

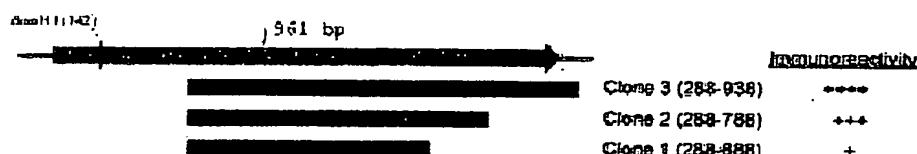
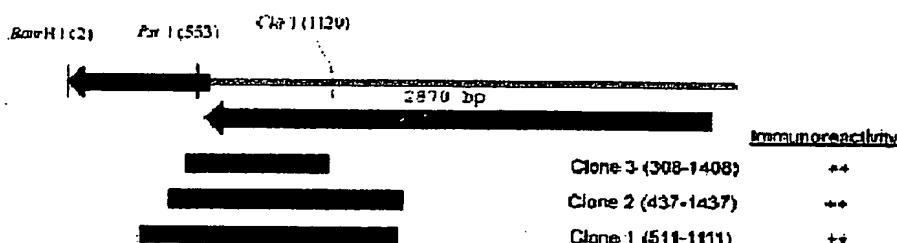
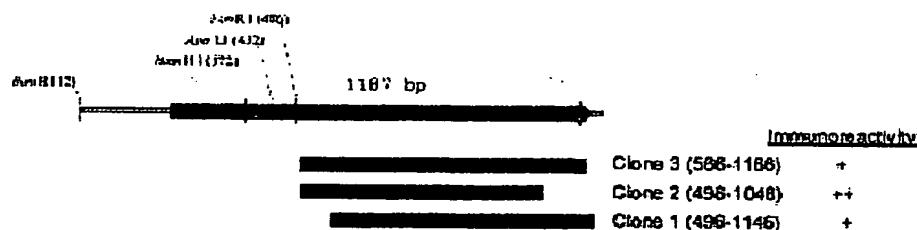
**A VV 39 kDa immunodominant antigen****B VV ATI in rpo35 region****C VV ps/hr**

Abbildung 4

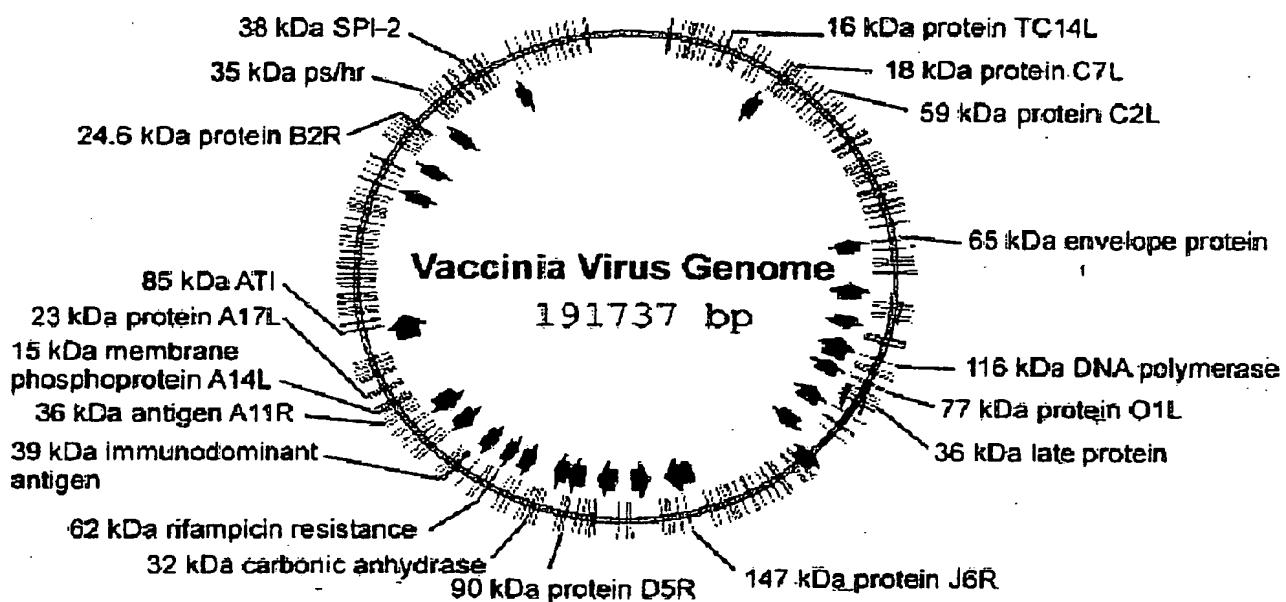


Abbildung 5

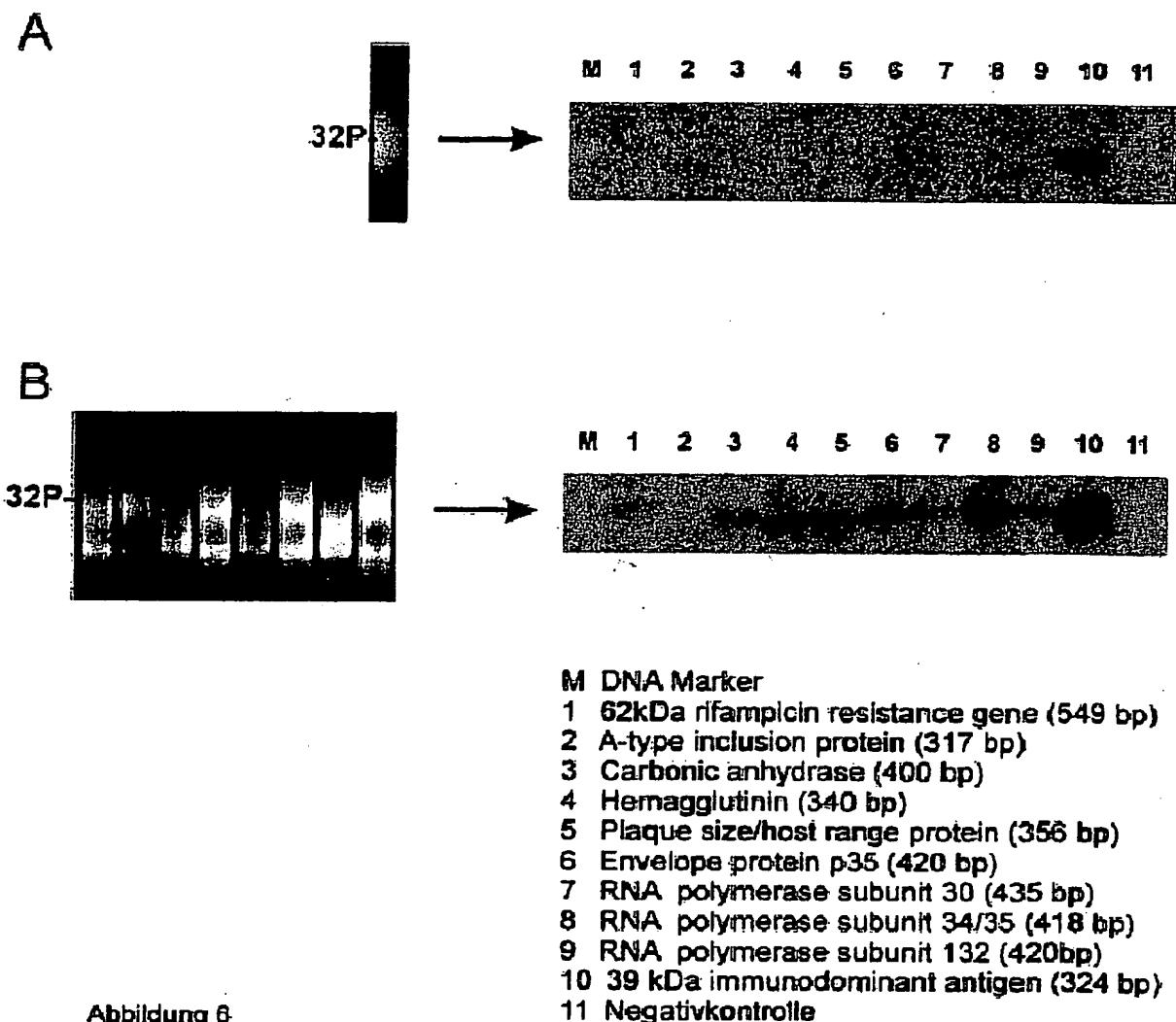
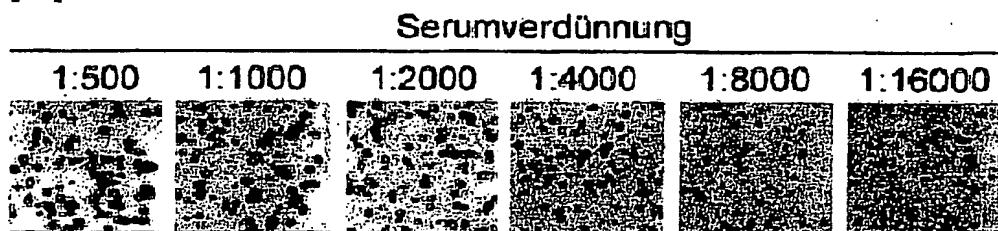


Abbildung 6

**A**

=> Anti-Vacciniavirus 39 kDa antigen Titer: >1:16.000

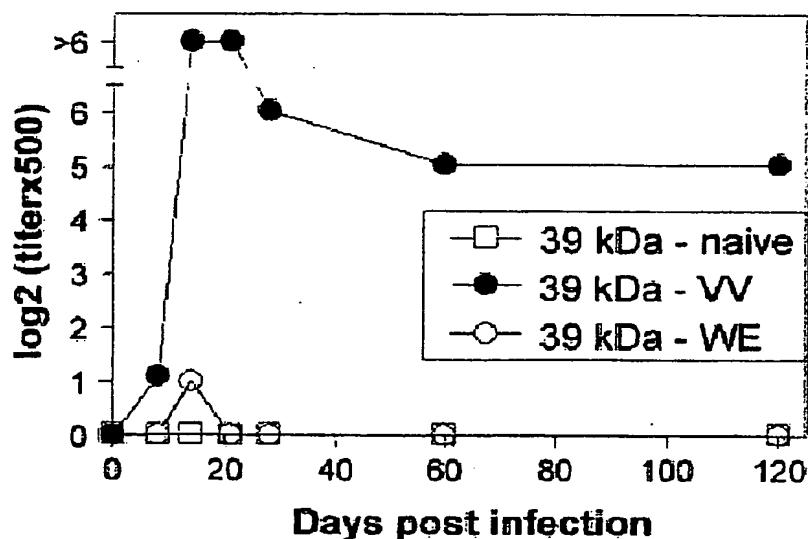
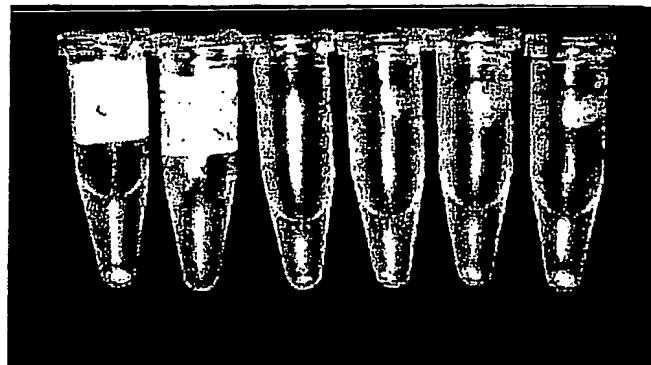
**B**

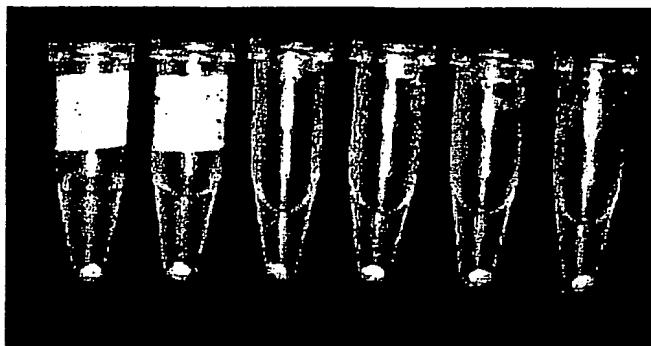
Abbildung 7

# SDS

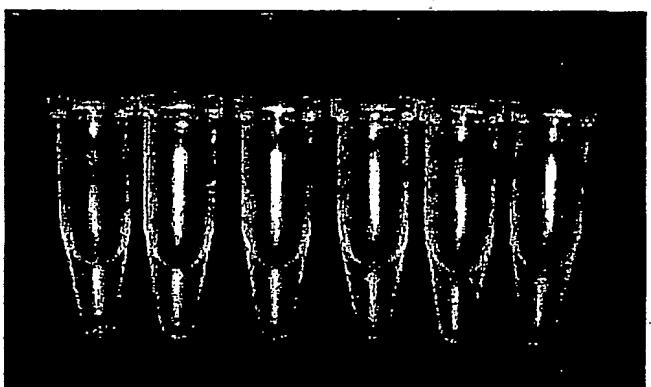
0% 1% 0,5% 0,2% 0,1% 0,05%



Gramnegative  
Bacteria



Grampositive  
Bacteria



Eucaryotic  
Cells

Abbildung 8

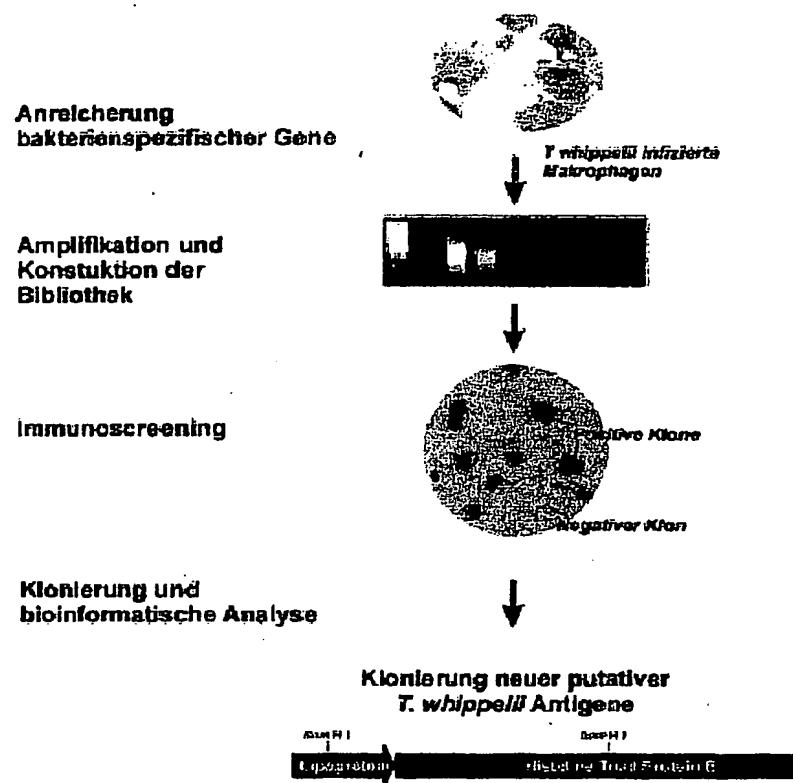
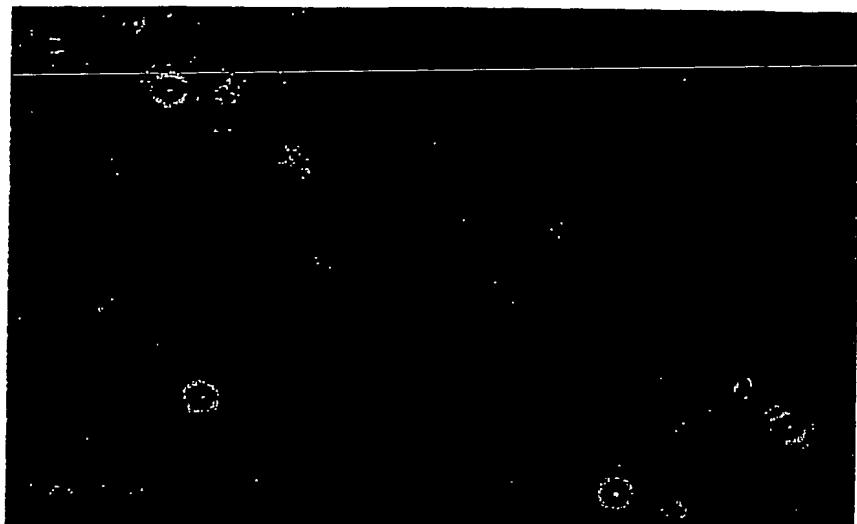
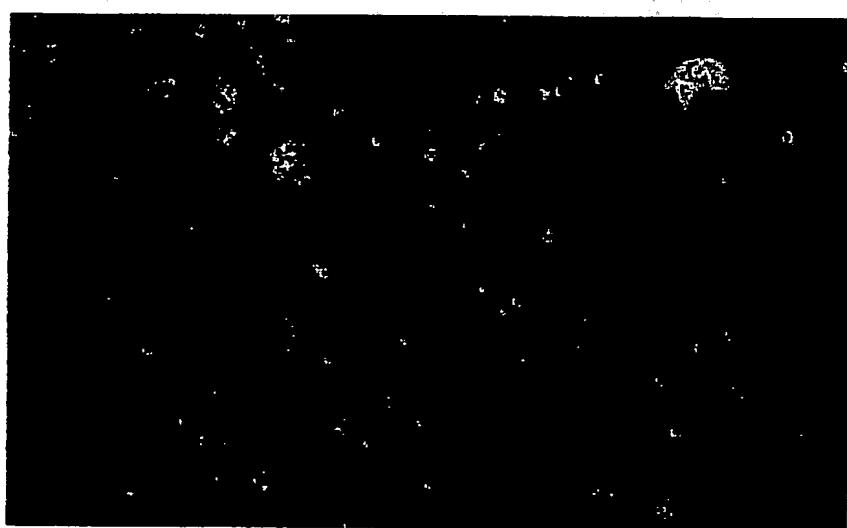


Abbildung 9

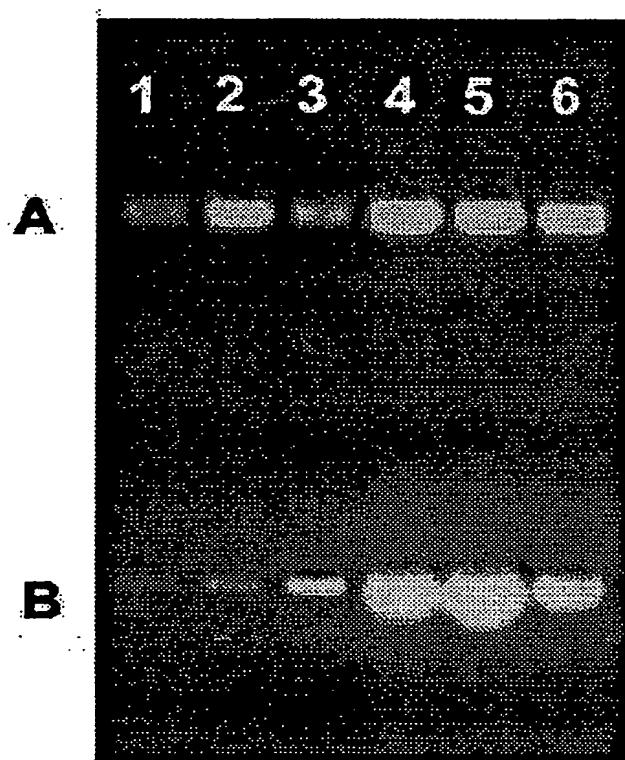
**A**



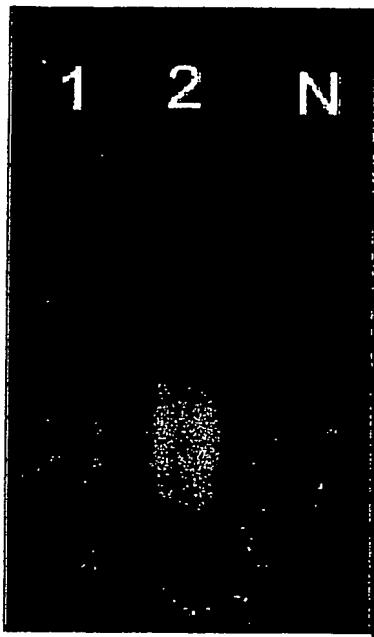
**B**



**Abbildung 10**



**Abbildung 11A**



**Abbildung 11B**

## Patientenmaterial



*Probenbehandlung bei  
virusen Pathogenen*



- milde Lyse
- PEG-Präzipitation
- DNase/RNase-Verdau



- DNA-Isolierung
- Amplifikation
- Herstellung Bibliothek



*Probenbehandlung bei  
bakteriellen Pathogenen*



- differenzielle Lyse
- differenzielle Zentrifugation
- Proteinase K-Verdau



- DNA-Isolierung
- Amplifikation
- Herstellung Bibliothek

## Abbildung 12

**Abbildung 13:**

ctaagacgtcgcatctctctgttccggcattggttcattattacgtctacagtcgtcaactgtcttcaagatctgatattcttagattggag  
ctgctaattctgttagcatttacggcattcactcagttgtcttcaagatctgagatttttagattggagtctgtaatctctgttaagatttcc  
ctccgcctcgatcgaggcgtcaacttattctctagttcttaatacgcgaacgcgtcatcaactcttgcgtgttccgttgctgt  
acattcatcgagtctagattcgagatcttaacgcgtcggttcccaagttctcggtactacagaaagcgtgtccctatcttgta  
tatttagcaattctgattctagagtactgatttgcttacgttagttactaatattgtcttggcatttatcaagatcccttgcatt  
tgatattccctacgaagtcggacagtcccatcgacattacgcgttacgttccatcgatcggagatcgtcatcggtttagccaca  
tagactgagttcaagttctcggttacaagatccatctactttccattcttaatagtatccagttccattcttagttctgaacgcattctcggt  
ccatcaagcgattctcaattctcggtatgttcttattcaatttctaataatctgaaccatcatgttccatttgaatattccctgttct  
tgatcttttgaagtcggcgttccgttataaacagaatcccttccaaagtcctaatcttactgagttatcactaagttctgcatt  
aattcggttagtttcttgcatttccaaactctgtttaactctccactatttccgatttccctcgatattctaaaccattcaattagtt  
taataacttagttgttaatcagcgttccatagccgttctgttaattgtggaaacataattaggacitcttaatggatigtatggcttgcata  
tcattttatcattattaggggatggacaaccttaatgggttgcctcatctccctccagtagcgtgtttctcaataccagtgttagtaat  
aggctttaggcaaatgttgcgtacgcgggacttccatccatcaagtattataatcgggttctacttcagaatattcttctaagagac  
gcgacttcgggagtttagagaagaactctgttctgtatctatcaacgcgttgcattcaacttcaagtttaaggatagcgaataccctcatcg  
atcatccgtatcttctgaaacaccatcatatgacatttcatgaagtctaacgtattgataaatagaatcagatttagtattaaacagatcc  
cttttagttaaaacgcataatgtatatttttagatctccagatttctataatatgatcacatgccttaatgtcagtgcattccatgatataatctgg  
caactatgggtgatgaaaaagataccggaccatatgcacgttgcataataactctgaaccactaagttagataatgatgtaatgg  
gagaaatattcagtatataggatgtcttggcgtcatatctgtactaaacacgcctaaacagttttgtatgtatcaatttccaatagattaa  
ttagagcagcaggaataccaacaaacatattaccacatccgtatttctatgaatatacatatcatgtttaaaaatctgtatagaagagcg  
aatatctcgctgacttaatgagtcgttagttcagcagcaacataactgttaatagaacataactttccgttagtattgattcttagact  
ccgcatcaacaccattattaaaatagtttatatacatcttataatctgcctccgttatactgcgtgaacgttctagttatacggaaacacttgc  
tttatctgttagttatgacttagttagatcacgaagaatattacgaattacatttctgttttcttgagagagatgattcagaactcaactcat  
gttccatagtttctacctcgtggcgaatcttggagttcgttgcattttcaataaggttctgtgacccat

35 kDa plaque size/host range protein (B5R) ORF 232 (167383-168336) +++ 7

ORF 232 (167383-168336) +++ 7

1

7

Ataaaaacgattccgttacgttgttatcgtaacctgcgtttattcaacatgtactgtacccactatgaataacgctaaattaac  
gtctaccgaaacatcgtaataataaccagaaagttagtacgtttacatgtatcaggatcatttcggatccaatgtctgcgaaac  
agataaaatggaaatacggaaaatccatgcaaaaaaaaatgtgcacagttctgattacatctgtaaactatataataaaccgctatacgaagt  
aattccaccatgacactaagttgcaacggcgaaacaaaatatttcgttgcaagaaaaaaatggaaatacttcttggaaatgatactttac  
gtgtccataatcgaaaaatgtcaacccttcattagaacacggatcgatcaccaggtaaagaaaaatactcattttggaaatataatgacta  
tcaactgtatgttgatgatggatgaggttattggcttcgtacataagttgtacagctaattttggaaatgttattccatcatgtcaacaaaaatgt  
gatataccgtctctatctaattggattaatttcggatctacattttctatcggtggcggttatacatcttagttgtaaaagtggtttataactaacc  
gatctccatcatccacatgtatcgacggtaaatggaaatcccgtactccaaatatgttacgaactaacgaagaatttgatccagtggt  
tggcccgacgatgagacagatttgagcaactctcgaaagacgttgtacaatatgaacaagaaaatagaatcgtagaagcaacttatca  
tataatcatatgtggcgtaacaattatggcgatattttatctccgttatactttagttgttccgtgacaaaaataatgaccaatataa  
gttccataatttgctaccgtaa

## 116 kDa DNA polymerase (E9L)

ORF 80 (59787-56767)

4

ttatgcttcgtaaaatgttaggtttgaaccaaacattttcaaaagaatgagatgcataaaacttattatccaatagattgactatttcggacg  
tcaatcgtaaaagttaacttcgtaaaatattttcgtactgcccggactttaaaacttcatacgataattgtctcatagttttaatatttacaagt  
tttggccatggcacattagccgacaaatataatgcaaaataatatcgtaatccaaagttctatgtttctggatttttattatattcagtaac  
caaatacatattagggttatcgcggatttataatttgagtgtactgcattcactcaacataataattctagaggagacgatctactatcaaatt  
tcggatcgtaaatctgtttctaaagaacggagaatatactatacatacctgattagaattcatccgccttcagacaacatctcagacagctg  
gtcttgtatgtcttaatcatattctatgaaacttggaaacatcttctagtttacttagtacccttattaaattcttcaggtacagatttgaatt  
gacgatgtgacttcatcggttatcttcgtattgcataatcagattttataccgcctcaaaactctattttaaaattttaaaca  
tactctattataatcagtcgttctaaactcttcgtattctatagacttacgtacatctgtctatctgttaaacacggagtcggatct  
catacagctacgaaaacgaaatctgtaatctataggcaacgcattttcacaatcggttataatctctatcgccatataaaatggattac  
ttaatggattggcaaccgttaacataccgttagataacttcgtccattttacttgcattttacttgcataatcagattttatccatggatgt  
gcaactcttagccgaaggctatgttagttagagcacttctaaatccatcagaccatataacttgcattttacttgcataatcagatttt  
catggaaatcatagatggccctttcagttgaactggtagcctgttttaacatctttatatctggctctctgcacaaaaatgtcttaatgtct  
ggaatggttccttctatcgatctatcgaaaattgtctttagatgagatgagggtcggtagtcttaggttccaaatgaaccgttaatctatc  
ggtagatatttctgaagcaatagctgatttttattttcttccaaatcttgcattttacttgcataacacggacttgcataatc  
ccaaagatacacacattaggatacagactgttataatcaaagattaatcattactaaacatttttttttttttttttttttttttttttttttt  
cttcataaggaaacttt  
ctctatattcgaataccatggattgaggaaggcacatgttgcgcacccgcgtctgttttttttttttttttttttttttttttttttttttt  
acacaaaacaagcatcatgaatacactatcttagccatctaaagctatgttttagattataatcttatacatctgagctaaatcaacgtcatc  
ctttccgaaaagataatttatgtatcattagttaaagtaggacataatagtcgactttaaatccatttccaaatatcttacgaatttttt  
catataatatcctcatcaacagtacacataattaccctgtggtaaaacccctttgcacatgcagcggcttgccttgcgtctgttagtacgtc  
cgataaaacgtcatcttcttaactcccttatttaatacttaccatgcacactgaacgcgttgcgttagatagaatccaatttttgc  
tttccagattt  
ttatgacttagactgatttcttcataaaataacagagatgtacagcttcccttttttttttttttttttttttttttttttttttttttt  
agacgatttagtaatatacttcgtacatcaaagttaatgtccgttaaaaggtaacgcgttagtgcacgcgttagttccaaacaatttgc  
acaaaactatttcagaacatagaacttagttctcggtatccattttccatt  
attttccctgtccgttaacatcttcattttatgtggcgtaaacaataatcggttaccactttaatcgatataacagtaacttgc  
atgtccgttagtt  
gttaataaaatacagaaggaaacttcttattcgaagtgtacacttctatcttagaaataatgtacgtatccatggatgt  
gttagtt  
gcaacagttacgtggatcgtcacaatgataacatccatt  
tatcccggttagaaatataaaaccaactaatattgagaattttcatccatgggttttttttttttttttttttttttttttttt  
atctgcgacggaggcattttctatctttaatatacttagattataacttattgtctcgtaatgtctatgttttttttttttttt  
aatggaggaggagacaatgactgatataatttcgtccgtactacgtaaataaaagtaatgagggaaatcgtaataatc  
gacatctggatttcagatataaaaatctgttttcccggtactttcaaccaattatgcacccgaacatccat

## **65 kDa envelope protein (F12L)**

**ORF 60 (43919-42012)**

2

ttataatttaccatctgactcatggattcattaatatctttataagagctactaacgtataattcttataactgaactgagatatacacccgg  
tctatggttccataatttagataaatgaatgctcgcaataactaatggcaaattgtatagaacaacgaaattatactagagttgttaaaggta  
atattttctatgagctgtccaataaaatttttgtgtactgcgttcaagtataatcatcttgatactatccagtaaaccgttttaagtctg  
aatattattatccatgtaaagccccataitcgactatcgaatatcctgtctgatagcagttcaatatcgacggacgtcaactgtata  
aaggtaggtatgtcatcatcgataaaactactggaatafggcgttagtaggtacggtaactttacacaacgcgatataactttcc  
ttgttaccattttaacgttagttggacgtcctgcagggtatgtttgaagaaatgatatcgagaacagatttgatcgatattgtggatc  
tgattatttactataatataatctagacagatagatgattcgataaataagagaaggatatactgttggtaggataatatacccccattccagtat  
ctcggatactcttataatgacacttagtaagaacatgtcttctattctagaaaacgaaaacatccatggactcattaaaacttctaacfct  
cctgattgtgtctcgaatgcctcgatcaaggattcaaggatgccatagattcttgaccaacgatttagatcgatctgtatttt  
attaaatcgaatggcgtctctgggttgcacccaaatgataacaatgtctgtaaagataaaaccgcaagaaaattatacgcatccat  
ccaaataaccctagcaccatcggtatataatgtattttatagattttccatccacaatttggccagttactgttagcaacggata  
tcgaatagattactcatgtaacctactagaatgatagttcgatctactgtcataatctttatccaaatctaagaaaattttaaatgat  
cactgttaaagttaaacaaaggattacccgggtacgtggatcatatatggattggccattatcatgtttatccataactgatacgg  
cgatggttttatatgtgtttgatctaacgaggaagaaattcgaccccacaattcatcttagatatgtttatataatcaaacggtacacatc  
aatttcgggacgcgtatgtttctaaattttatccaaatataatgtatccatatgtttatccatactgtcaactatagttacacctaga  
gaacttacgatacatctgtttctataatgttaaattttacaaatctataacatgctaaaccccttgcgacaaccattcattatctgatag  
gaatctgtattctcgataccgttctaaagccagtgcataatctccctgttgcgtggaaacgcgttgcataatatcgatcaacggataat  
ctgaagttttggagaataatgactcatgatctttcgatccataaacaatctagacatagatggaggcgatgtttatgtca  
atgagtgtcaatccataacttctaatcttgtaatattcatcgacataactatctatgttgcataatgttgcattttatgtca  
catttcgtgccaatataatgtatccatgtttatgttgcataatctcaatgttgcattttatgtcaagattgttaccctgtt  
aacat

62 kDa rifampicin resistance gene (D13L)

ORF 145 (113026-111371)

+++

1

ttagttattatctccataatctggtaatacttacccctgatcgtaagataccttatacaggcattacataactaccaattgttttgtaca  
 aatagattggatgggtgacatccatggtaataactactcgAACAGATAGTTTATCTTCCCTAGATACTGGCGTAATAGTTGTC  
 GCCTAAAGAATATCTTGGTAAAGTAAAGTAGGGTCTTCATTGCTTTGTCAGTAGTCATTATAAATCTCAGAGATGGTC  
 GTTCTCTGAATATAGAACATCATTCCAATCTAACTCTAGTCTAGAAATAATCGGTCTTATTCTAAAATCTATTCCCTGATGAAGGGATC  
 GTTAATGAACAAATCCTGGCCTTGTTCGGCTGATCTTATTCTCGTATAGACGTTGACTAGTCACAGACTACAGGAATAGAT  
 GTATCGATGATGTTGATACTATGTGATATGTGAGCAAAGATGTTCTCTAGTGGCATCACTATATGTCAGTAATGGCGGAAAACCTTTAG  
 AATGTTATATATAAAGAATTTCGTTCCAACATTAGCAGATTAGTGAAGATAAACACTCATATTATCGGAACATTATCAATTTCAC  
 ATACACATCAGCATCTGAATAGAACGATAACCCTCTGGAACCTCTACGATCTCGGCAGACTCCGGATAACCAGTCGGTGGGCCATC  
 ACTAACAAATAACTAGATCATCCAACAATCTACTCACATATGCTATATAATCTTTCTATCTGTGAGTACCTGGATACGAAATAATTATT  
 ATCCGTTTCCATAATAAGGTTAGTATAAACAGAGAGCAGTTGCGCATGAACCTCAGTACAGTCGCCGTTGGTTATTGACCT  
 ATTACTCTCTAGGTTCTATAAACGATGGTTAATTGTCATTCTTAACCATAATCCAATAAGCTCAATTAGGAACATAAACAAATTCT  
 TGTGAACGTTCAAAGTCGAACGAAGAGTCACGAATAACGATACTGGATACTGGATTACCGTTACGTAATTGGATACGGAT  
 AGTTAAGACTGCTGAATGTATCTCCACATCAAACGGAGTTAATATAAACGTTACTGTGAGTGGTTAATAGTGTATTAGGAGTTAG  
 GCCAATAGAAATACTATCATTAGTTCTAGAATACTCCAGAGTGTTCAGAAGCAATTGTTATTGATAACAATTATATAATTCTCGCCCTCAAT  
 TTCCCAAATAACACCGTTACACGAAAGAGATACTGTTGATAACATATCCAAACATATGGTACGTAACCGATACTTCCCATACCTT  
 AACCTCTGGAAAGTTCCAACCTCAGAACCAAATGATTAAGCGCAGTAATACTGATCCCTAATTGCAAGCTAGCGATAGCCTGATTGTCTG  
 GACCATCGTTGTCATAACTCCGGATAGAGAAATACTGGCATATAAAAGTTGAAATTGACTATCGACTGCGAAGACATTAGACCGT  
 TTAATAGAGTCATCCCCACCGATCAAAGAAATTAGTATGATGTTTCTAT

32 kDa carbonic anhydrase-like protein (D8L)      ORF 137 (107120-106206)      +++      1

ctagtttgttttcgcgaatatcgactcataagaaagagaatagcggttaagtataaacacgaaataactatggcaataattgcgaatgt  
 ttatcccttcgatataattttgataatatgaaaaacatgtctctcaaatcgacaaccatctataaaatagtctcgcgctggagag  
 gtagttgcgtcgataatctccccagaataatatacttgctgtcgctcaatttatacggtttctatagttctgttatataatacggt  
 ttccatcatgattagacgacgacaatagtgttcaaatttagatagttgatcagaatgaatgtttatggcggtggaaaaattatccatacagc  
 gtctcgagagtgtgtatagttgttcctagatatgtaaaataatccaacgtactaggtagcaaaattgtctagataaaatactgaatcaaacg  
 ggcgcagacgtattagcggtataatggaaatccaattgattgactatctttgaaaatacattttatgatccgatacttgtaaatata  
 aataatgataatccatcatcggtttttgccttcataagaactatattttcttattccatgaacaagattaatctccagagtatttgt  
 cacatctatcaagtgtttggatccataatcgcttccttcccaatataatgttagtgatgataacacatattcattggggagaaaccctcca  
 ctatataatccctttaaaattatccctactagtttccagtggtctggatagtgggtgttcgactcattataatgtatgtctaaacggcttca  
 tcgcgcgttagaaattgtttttagttctatattaaataggagatagttgtgcggcat

36 kDa late protein (I1L)      ORF 87 (63935-62997)      +++      1

ttattcagcattacttgatataaataggcacagtcaaaccactctcgatacattaactctctattttttaacaaattctgc  
 atatcttcgtaaaagattcttgaacttttagatatctatcgactcttagatgaaatagcgttcgtcaacatactatgtttgtatacataaaag  
 gcgcccatttaacagttcttagtgacaaaatgctagcgatccatggatcctttagaatcacatagattgacgattcgtctcttagtaactc  
 tagtaaaataatcatacatctatcgcaaaataatattatccgtacttgaggagatctaaacaatctatgtttgagaacatcgataagttc  
 atcgggaatgacacataactatcttaatagaactcttcatccagttaatggattcgtccatccaactgattatgagatcttctatTTT  
 atcatttccagatgatgtatgtccattaaaggtaattgttagcgctcttttagtctagcagccaatactttaacatcataatatcgat  
 tacaaggagatgatattatctatggtattaagaattcgtttcgacatccgtcaaaaccaattcctttgcctgtatcatccagtttccatcc  
 ttgtaaagaattattttctactagactattaaagactgataaggattccatcaattgcacaatccaaacttttcaaaaaactagacttta  
 cgagatctacaggaatcgctacttcaggttcttagcttgtgattttcttgcggacatttctgtgaccaactcatctaccatttcatgt  
 ttagcagtggaaataagcttcaatgcacggcactgatactattgaaaacaggttgcattcaaaattccgccccat

## 16 kDa protein (TC14L)

## ORF10 (10995-10567)

+++ 1

Ttaattgcaaaagatctatataatcattatagcgttgacttatggactctggaaatcttagacgtacatctataatcatggcatattta  
 tacattgtttatagcatagtagttatctacgtatgttagatattctctcaatgaatcaatcacataatctaattgttaggtttatgcataatgcat  
 ttccagcaggtaatgttcttagattcgttgatggcaatggctatacatgtataatccgttattgtatctaattgttagacatctgaaccggattctag  
 agtaaagatactagagattgtttatattatctaacagccttgcagaagatgtttctccctcgtttgcataatcatgtttaagataaggt  
 aggcaaatgtttatagtagactaagaattggcaagcataagacat

## 38 kDa serine protease inhibitor 2 (B13R)

## ORF 421 (172562-172912)

++ 4

Atggatatctcaggaaatcgcatctctatgaaaggagagaatgtattcatttccagcgtcaatctcgtagtttgcataactgttat  
 tatggagctaattggatccactgctgaacagctatcaaaatgttagaaaaggaggagaacatggataaggtagcgctcaaaatatctc  
 attcaaattccataaataaaagtatatggcgatattcgtccgtttaagattccctttgagaaaaattggcgataagttcaactgttgact  
 tcactgattgtcgactatagatgcaatcaacaagtgttagatattttactgagggaaaatcaatccactattggatgaaccattgtctc  
 ctgataccctgtctccatgcaattttactgtccgtatactttaaagcaaaatggtagtgcgttcccttgcggcatttcgaaaaaggaaatttacc  
 agtgcattttccatgatgttt  
 cgtatctccgacggaaatggtagatgttagatgtatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttccat  
 ttttcaatcatagaactgccccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttccat  
 ctaacagatacaaatttaagaaatggtagtactctctggaaagctacgtttatcgatgttccatgttccatgttccatgttccatgttccat  
 aatctggtagtactcttagtaaagtctcgaggactgacagaggtgttcgggtcaactggagattatgcataatgttcaatcgatgttccat  
 cgacgctatgtccacaaaacgtatataatgtcaatgaagagtatacagaaggcgtcagcaacttgcactgggtcagactgt  
 catcaacaatttacaaatggtagtgcattatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttccat  
 ctccgacaactaattgttca

## 18 kDa protein (C7L)

## ORF 24 (19257-18805)

++ 3

Ttaatccatggactcataatcttatacgggattaacggatgttctatatacggggatgagtagttctcttttaactttataacttttactaat  
 atattttagactgtatgggttaatagtgtttgaagagagctcgftctcatcatcagaataaaatcaatctctgttttttttttttttttttttttt  
 cagccctatataattacgttaatagaacgtgtcatctaccttattaaactttcaccgcatgttgcataatcggtaatccttgcacccgtcg  
 atttccgaccaatctggcgatataatgaatctaaactttaatttcttgcataattcgaataatttttagttgcattccgtatgttgcattcccttt  
 taactgttaatttctcaacgcgatctccatataatgttgcattccgtatgttgcattccgtatgttgcattccgtatgttgcattccgtatgttgc  
 ctccgacaactaattgttca

### 24.6 kDa protein (B2R)

ORF 226 (163876-164535) ++ 2

Atggcgtatttacgcacacgctctcggtggtacgacgagaatcttcatgccttcgtggaaatatcatcgacttgtgccaatgatgtca  
ggaaatattctgttgtcagttataataacaagtatgacattgtaaaagacaaatataatgtggtgttacagtcaggtaacaagagagatata  
ttggagcactgctgcctatgttgagtgcaatgaatatctacaaattggagatccgatccatgatcaagaaggaaatcaaattcttatcatc  
acatatcgccacaaaaactactatgtctctaagcggaatcgggtacgagagtctagacttgttgttggaaaggaggtagggattcatcatc  
gtacttggaaacagggaaacgctgtatatggaaaagttaacatgattttctactatcaaagagaaggccaaagaaatgaggactacacttagc  
caggacctataattgattaccacgtctggataggagattgtatctgtcaagttactgcgtggacgtacatggaaaggaaattatgagaat  
gagattcaaaaagggtgcggtgcgtccgatccaaatctggtaaaagttaacttggggagaatgatacagaaaatcttcttactatata  
cgccggcaccatcgaggtaa

### 36 kDa protein (A11R)

ORF 164 (124976-125932) ++ 1

### 15 kDa membrane phosphoprotein (A14L)

ORF 167 (126785-127128) ++ 1

Ttagttcatggaaatatcgctatgattgtatgaatgactccgctaactctgtggggcgcagtgccttccccacatagaataaatttagca  
ttccgactgtgataataatccaagtataaacgccataactcaataactttccatgtacgagtggactggtagacttaactaaagtcaata  
aaggcgaagatacacgaaagaatcaaaagaatgattccagcattagcacgccccaaaaataattccaatcataagcatcatgtccat

### 147 kDa protein (J6R)

ORF 117 (86510-90370) # 1



77 kDa protein (O1L))

ORF 84 (62477-60477)

++ 1

ttaacgaggttccatttatcatccaatattattgtaaatgacgttgatggacagatgataacaataagaaggtaggtacggtaccttgtccaccat  
 ctccctccaattcatgctctatttgtcattaacttaatgtatgaaaacagtgccacatgctccatgacagtgtgtacactttggatacaa  
 aatgtttgacattgtataattgttaaagactgtcaatctataatagatagtagctataatattctatgtatggattgaagaagatgacaatct  
 tggcatattgtatcattiaacacagacatggtatcaacagatagcttgaatgaaagagaatcagtaatttgaataaagcgcttcgtatgga  
 gtgtccgtataccaacatgtctgatatttgatgtattccattaaattttatgtttttctcgtaaacagcattctgtcaacggacc  
 ccaacatcggtgaccgattaagtttgattgtttccgttaaggcgatctagtcagatcgatgcctatccaataatccatcatctgtgc  
 gtagatcacatcgtaacttttaattctctatagaagagcgacagacatctggacaattacagacagcaattctttattcttacagatgt  
 aagataacttgaagacattctatgtatgtcagaattttggataacacgttattgtatggatctgttaccataattccttgcattttgatggctgataatgt  
 gtcagagcacaagattccaatcttgcacaattttgaccattatctttgtttgatatctatcatcagacagcatggcgatcgacaacaca  
 gggattaagacggaaagatgattctcaacatctcaatggataccctgtattttctggcattatctatgtgcgagaatatcc  
 cttagagaatcagtatccttttgcattgtatgttgcattcaatgacatggacgtctaaaccccttattctatcaccagattgcattttgcattttgt  
 ctctttctttatcataatgtatctaaattcatcgccaaatgtctatctaaatcataatatgagatgttaccctctacaaatatctgtc  
 tccaaatgttagatctacatcagtttgcattccaaattaaacatggcaacggatttaattttatattcccttattaaagtccctgtcgataataa  
 cagaatgttagataatcatttaatccatcgatcggttggaaagatgctgttgcacaaaatctttaatgtctgtatgaaggtagggactatatact  
 aacatcttgcattaaataaaattataacattgtccataggatactttgtacttagtttatacacatcttcatcggtatgttgcattttgcatttt  
 tgaacagagggttatattattcatcagatatacgaagaacaatgtccaaatctatattgttaatattatattatagatgtatgttagctcctac  
 aggaatatctttaactaagtcaatgattcatcaaccgttagatctattttaaagttaatcatataggcattgattttaaaaggtagtgatgcctt  
 actacattcttcatattaccattccaagtcaactgtgtgtatggaaagattatattctatcataagcttgcattttggccccgataccattaa  
 gaattcttgcattataaggaaacagcttttaggtactcatctacttcaagaattttggagagccttaacgatcatgtgcattttatttt  
 aggagggaaaaacctaacattgagaatgtcgagttatagcttccagatacgtgattttggcaatagtccgtgtatccataatccagt  
 aacacgagctggcttgcttagacacccttcaatgttaattttgaaataagcttgcattttgataaagccttcgtcaaaatccggatacatga  
 cat

59 kDa protein (C2L)

ORF30 (24156-22618)

+ 4

ctattgtagaaattgtttcacagttgctaaaaacgatggcagtgacttatgagttacgttacactttggagtcacatcttagtaaacatat  
ataatattcgatattacgagttgacatcatcgaaacaatccaagtatttgatttggataatattcgatttgcattgtctataattaagatataa  
caccgcaagaacacacgaacatcttcctacatggtaaagtacatgtacaattctatccatttgcatttccttaactatataatttgtatagataat  
tacgagttcgttgagtaattccagtaattacatagatgtcgccgtactctacagcataaaactatactatgatgtctaggcatggagac  
tttttatccaacgatttttagtgaacacattccacatcgtaataactacatattttcatacgtggataaaactccacccattacatataatcatc  
gtttacgaataccgacgcgcctgaatatctaggagtaattagtttggaaagtcttattccatttcgaagtgccgtgttcaaataattctgccac  
acccgttigaaatagaaaattctaattccctattacatataactttccatgttaacacaagtaacttctgattttaacgacgacatattag  
taaccgtttccattttcgttcaagatctacccgcatacggaaaaacatgtctattgttaatcatgccccaataatgtatagacaatta  
gtaaaacatttgattatagaattgtctatctgtattaccgtactatcgccatattctgtccttaggagagtaatgggtattgtggatataa  
cagagtttaatgactactatattatgtttataccatttcgtgtcactggctttagattttggatatagttaatccaaacaatgatatacgatt  
cgcatagtattagtctaaaacttggatgtaaaatgttgcattatctacatcggtttggatttatgtatccactttaataatcatagctgt  
catcctcatgatttacgttaacgtttcggtggataagatagttgtcagttcatccttigataattttccaaattctggatcgatgtcaccgca  
gtaatattgttgcattttctgacatcgacgcattatagtttttaattccatatttttagaaaagttaaacatccattacaatttgcatt  
atattatgaatcatagttttacacatagatctactacaggcggaaacatcaatttacggcagcaacttagtattttctacattgtttatgg  
gtatgtttatcttccagcgcatatgtctaatagcgttcaaaacgcgtatgtttataccattcaatataatcgcttcatccatttagatgg  
atccctgaatgcgtttaaaaaattatacggagacgcccgttaataattccatttacttgtataattccccattgtatagaaaatcagcatt  
ccat

90 kDa protein (D5R)

ORF132 (101420-103777)

+

1

atggatgcggctattagaggtaatgttatcttgcctaagactatagggtcccatcagcatgtagacaaaatgaagatccaaggatt  
 cgtagaagcatttaaatgcgacgagttaaaaagatataattgataataatccagaatgtacactattcgaaagtcttagggatgaggaagca  
 tactctatagtcagaatttcatggatgttagatttagcgcgtctagacgaaatagattttacggctattcaagattttattatcgaggt  
 gtcaaactgttagctagattcgcgttacagaatgcggccattcatgaaaatgtataaaaatccatgagatctaattttcattgactaa  
 tctacaaaatagagataaaacaagtttcalattatcttttagacacgtataccactatggatacttgtatgtacgtacaaacactattag  
 aattaagttagatcatctgaaaatccactaacaagatcgatagacactgccgtatataaggagaaaaacaactctcgggtttaggtactag  
 gaaaaatccaaattgcgacactattcatgtaatgcaaccaccgcataatataagattacctattacgtggatataacaaca  
 atagttattactttctacaacgcgattggaggatttagttctgataagttatggaaaccagggttattcattcgaagacgtataaa  
 agagttcaaaaatattcattaattctataataaaactttatgtatctcgatgaaaataatttacaacggtaccactggtcatagattacgtaac  
 accttgtcattatgtaaaaaacgatcgatcaaaccatccgcataactatcggtggaaaatggtctattagaattttacaaaactggtaatc  
 cacatagttgtaaagtaaaattgtccgtggatgtaataaaactgtttatattgcacaaagaattttagacactaactctgttttattaacc  
 aacgaggagactatatagtttgattaataattcatggaaatttacagcgaagaacccttgataacaaaactaattctgtcaataagacat  
 caactacctaaggaatattcaagcgaattactctgtccgaggaaacgaaagactgttagaagactaaccatacgagacatgttagattca  
 gtagagaccgataccatccggataaaacttccgtttaaaatgggtattggacctgttagacggaaatgtttactctggagatgtctaa  
 aaaatatacgtgtactgtatcaaccggatttaatttacgatacaaagtctgtcgaagacagtccagaaatgaaagatgttagatgaaat  
 attaacgatatccaaccattaacggatgaaaataagaaaaatagagactatgtatgaaaaacattatctagttgtttatgtggctaccaa  
 aggatgttaacatttttggagaaactgcaactggaaagtgcacaacccaaacgtttttaaagtctgtctatcggtgaccgtttgt  
 gaggggtcaacaatttacagatgtattggataaaggacctaattccattatcgtaacatgcattgtaaaagatctgtattctgttagcga  
 actacctgtttgcctgttagggatcaaagaaaattagatctgataatattaaaaatgttagacagaacctgtgtcattgaaagaccgttt  
 ctccaaattaaataatagaaaccatgtcataattcattatcgatactaattacaaacctgtctttaggatagataacgcattaaatgag  
 aagaattggcgtcgattcagaacacacacttttcaacccctgttagagaggctgctgaaaataatgacgcgtacgataaaagtcaa  
 actatttagacgaggggttagatgtaaaatacaaaataatagatatagtttgcatttctatacttgtgtgtaaatggtacaaaaatata  
 gttcctattatgaaactatatacctacaccgaaagagattcctgacttctatctcaaaaataggactctgttttagtctatgtctgt  
 gcatattccattaaatgacggaccctccaaaaaggatataattgtacgataatgtggttactctccgttactttccaaacagaaaata  
 tccaagtatttaattctagactattggacacgatataagagacttcatcaatagacataagaaatttgcctatgtgatgaatatctg  
 aatataatattcatagaggatatttcatctccgtaa

## 23 kDa protein (A17L)

## ORF170 (129314-128703)

+ 1

ttaataatcgtcagtatttaactgttaatgtggtatatacattcaccttattccgcagtataagggtttgcaggtatactgttcagg  
 aatggttacattatacttctatagtcctgtcttcgttcatcacatgtcaaaagaacagaataaaacaaaataatgtaaagaaaataat  
 aaatatctgtgaattcgtaatacattgtattgcataataattacagcagctacaataccacacaatagacattcccacagtgtgtccattacc  
 tccacgatacatttgagttactaagcaataggtataactaagcttagtaagaggcaatagaaaagatgagataaaatcatcaatata  
 attagaggagggctatatagagccaagacgaacaaaatcaaaaccgagtaacgttctacatcatttttgaagattccaaataatcatt  
 cattccctccataatcggttgcattcataccctccatctttaggcataaaacgattgtctgtttctgtaaataatcttcatcaagcactccag  
 acccgcgagagaagtgcgtcaagcatattgtataatcttaataactcat

## SEQUENZPROTOKOLL

<110> Ugur Sahin, Özlem Türeci, Burkhard Ludewig

5 <120> Verfahren zur Identifizierung biologisch aktiver Strukturen mikrobieller Erreger

<130> 268-3 PCT

0 <150> DE 10108626.1  
<151> 2001-02-22

<160> 22

5 <170> PatentIn Ver. 2.1

<210> 1

<211> 32

<212> DNA

0 <213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
Adaptor-N(6)-Oligonukleotide

5 <400> 1

gatgttaatac gaattggact catatannnn nn

32

0 <210> 2

<211> 25

<212> DNA

<213> Künstliche Sequenz

5 <220>

<223> Beschreibung der künstlichen Sequenz:  
EcoRI-Adaptor-Oligonukleotide

0 <400> 2

gatgttaatac gaattcgact catat

25

0 <210> 3

<211> 24

5 <212> DNA

<213> Künstliche Sequenz

0 <220>

<223> Beschreibung der künstlichen Sequenz:  
MfeI-Adaptor-Oligonukleotide

0 <400> 3

gatgttaatac aattggactc atat

24

5 <210> 4

<211> 846

<212> DNA

0 <213> Vaccinia virus

0 <400> 4

	ttacttttgg	aatcgtaaa	aacctttgac	tagttgtaga	atttgatcta	ttgcgcctacg	60
	cgtatactcc	cttgcatacat	atacgttcgt	caccagatcg	tttggttcgg	cctgaagttg	120
	gtgcataatct	tttcaacac	tcgacatgag	atcccttaagg	gccatatcg	ctagattttg	180
	tttagatgtct	gctctggat	ttggattttg	ttgtgtctgtt	gtacatactg	taccaccagt	240
5	agggtgttagg	gtacatacacag	tggccacaat	aggaggttga	ggaggtgtaa	ccgttggagt	300
	agtacaagaa	atatttccat	ccgattttg	tgtacatgtt	gttggggta	acgtctgaga	360
	agggtgggtt	gatggcgccg	tcgtcggtt	ttgatctta	ttaaattttag	agataaatatc	420
	ctgaacacgca	ttgctcgccg	tcaacgcgtt	aaggagtgaa	ctcgccggcg	catcgtatc	480
	ttcagacacgc	caataaaaaa	gattagacat	atcagatgtat	gtatttagttt	gttgcgttgg	540
10	ttttgggtgtt	ggagcagtac	tacttaggtt	aagaatagga	gccgggtttag	ctgttggAAC	600
	cggctgtgga	gttatatgaa	tagttgggtt	tagcggttgg	ataggctgtc	tgctggcgcc	660
	catcatattt	tctctagcta	gttgttctcg	caactgtctt	tgataatacg	actcttgaga	720
	cttttagtctt	atttcaatcg	cttcatcctt	tttcgtatcc	ggatccctttt	cttcagaata	780
	atagattgac	gactttggtg	tagaggattt	tgccagcccc	tgtgagaact	tgtaaaagaa	840
15	gtccat						846

<210> 5  
<211> 2178  
<212> DNA  
<213> Vaccinia virus

	<400>	5					
25	ctaagacgtc	gcatctctc	ctgtttcgcc	attgggttca	ttattacgtc	tacagtgcgt	60
	caactgtctt	tcaagatctg	atattctaga	ttggagtc	ctaattctc	tagcatttc	120
	acggcattca	ctcagttgtc	tttcaagatc	ttagattta	gattggagtc	tgctaatttc	180
	tgttaagattt	cctccctccgc	tctcgatgca	gtcggtaaac	ttattctca	gttctcta	240
30	acgcgaacgc	agtgcataaa	tttcttcgt	gtttcctgg	ttgcgtgtac	attcatcgag	300
	tctagattcg	agatctctaa	cgcgtcg	ttttcctca	agttctc	gtactacaga	360
	aaggctgtcc	ctatcttgtt	gatatttagc	aatttctgat	tctagagtac	tgattttgt	420
	tacgtagtta	ctaataatttg	tcttggcctt	atcaagatcc	tccttgtatt	tgtcgcattc	480
	cttgatatacc	ctacaaggtc	tggacagttc	ccattcgaca	ttacgacgtt	tatcgatttc	540
	agctcggaga	tcgtcatcgc	gttgttttag	ccacatacga	ctgagttcaa	gttctcggtt	600
35	acaagatcca	tctacttttc	cattccta	atgtatccagt	tcctttctca	gttctgaacg	660
	catttctcg	tcccttatcaa	gcatattctct	caattctcg	atagtttct	tatcaatttc	720
	taataaaatct	gaaccatcat	ctgtccccatt	ttgaatatacc	ctgtgttctt	tgatctctt	780
	tgttaagtccg	tcgattcttt	cgggtttata	aacagaatcc	ctttccaaag	tcctaatctt	840
	actgagtttta	tcactaagtt	ctgcattcaa	ttcggtgagt	tttcttctgg	cttcttccaa	900
40	ctctgttttta	aactctccac	tattnccgca	ttttcctcg	catttatct	accattcaat	960
	tagtttata	ataactagtt	ggtaatcage	gattcctata	gccgttctt	taattgtggg	1020
	aacataatta	ggatcttctaa	atggattgtt	tggcttgata	gcatcatctt	tatcattatt	1080
	agggggatgg	acaaccttaa	ttgggttggc	ctcatctcct	ccagtagcgt	gtggttctt	1140
	aataccagt	tttagtaatag	gtttaggcaa	atgcttgc	tacgcggca	cttcctcatc	1200
45	catcaagtat	ttataatcgg	gttctacttc	agaatattct	tttcttaagag	acgcgcactt	1260
	gggagttagt	agaagaactc	tgtttctgtt	tctatcaacg	ctggaatcaa	tactcaagtt	1320
	aaggatagcg	aatacctcat	cgtcatcatc	cgtatcttct	gaaacaccat	catacatgacat	1380
	ttcatgaagt	ctaacgttatt	gataaaataga	atcagatttta	gtattaaaca	gatccttaac	1440
	cttttttagta	aacgcataat	tatatttttag	atctccagat	ttcataat	gatcacatgc	1500
50	cttaaatgtc	agtgttcca	tgtatataatc	tgaaacacta	atgggtgtat	aaaaagatac	1560
	cggaccatata	gctacgttga	taaataactc	tgaaccacta	atgatataat	gattaatgtt	1620
	aaggaagagg	aaatattcag	tatataggta	tgtcttggcg	tcatatctt	tactaaacac	1680
	gctaaacagt	ttgttaatgt	gatcaatttc	caatagatta	attagagcag	caggaatacc	1740
	aacaaaacata	ttaccacatc	cgtatcttct	atgaatatac	catatcatgt	taaaaaatct	1800
55	tgtatagaaga	gcgaatatct	cgtctgactt	aatagagtgc	agttcagcag	caacataagt	1860
	cataactgt	aatagaacat	actttccgt	atgattgtt	ctagactcc	catcaacacc	1920
	attattaaaa	atagttttat	atacatctt	aatctgc	ccgttaatcg	tcgaacgtt	1980
	tagtatacgg	aaacactttt	atttcttatac	tgttagttat	gacttagtga	tatcacaag	2040
	aatattacga	attacatttc	ttgtttttct	tgagagacat	gattcagaac	tcaactcatc	2100
60	gttccatagt	ttttctacat	cagtggcgaa	atctttggag	tgcttggta	attttcaat	2160
	aaggttcgt	acctccat					2178

<210> 6  
<211> 954  
<212> DNA  
<213> *Vaccinia virus*

25  
<210> 7  
<211> 3021  
<212> DNA  
<213> Vaccinia virus  
30  
<400> 7  
ttatgcttcg taaaatgtag gtttgaacc aaacattctt tcaaagaatg agatgcataa 60  
aactttatta tccaatagat tgactattc ggacgtcaat cgtttaagt aaacttcgta 120  
aaatattctt tgatcactgc cgagttaaa acttctatcg ataattgtct catatgttt 180  
aatatttaca agtttttgg tccatggtag attagccgga caaatatag caaaataata 240  
tcgttctcca agttctatag tttctggatt atttttatta tattcgtaa ccaaataacat 300  
attagggtta tctgcggatt tataatttga gtatgcatt cgactcaaca taaaataattc 360  
tagaggagac gatctactat caaattcggg tcgtaaatct gtttctaaag aacggagaat 420  
atctatacat acctgatttag aattcatccg tccttcagac aacatctcag acagtcttgt 480  
cttgttatgtc ttaatcatat tcttatggaa ctggaaaca tctttcttag tttcactagt 540  
acctttatta attctctcag gtacagattt tgaattcgcac gatgctgagt atttcatcgt 600  
tgtatatttc ttcttcgatt gcataatcag attcttataat accgcctcaa actctatttt 660  
aaaatttatta aacaataactc tatttataat cagtcgttct aactcttctg ctatttctat 720  
agacttatacg acatcttgac tgtctatctc tgtaaaacacg gagtcgttat ctccatcac 780  
gtacgaaaaa cgaaatctgt aatctataagg caacgtatgt ttcacaatcg gattaatatc 840  
tctatctgtcc atataaaaatg gattacttaa tggattggca aaccgtaaaca taccgttaga 900  
taactctgtc ccatttagta ccgattctag atacaagatc attctacgtc ctatggatgt 960  
gcaactctta gccgaagcgt atgagttatag agcactatctt ctaaattccca tcagaccata 1020  
tactgagttg gctactatct tgtacgtata ttgcatggaa tcatagatgg cttttcagt 1080  
tgaactggta gcctgttta acatctttt atatctggct ctctctgcca aaaatgttct 1140  
taatagtcta ggaatggttc cttctatcga tctatcgaaa attgctattt cagagatgag 1200  
gtcggtagt ctaggttac aatgaaccgt aatatatcta ggaggtggat atttctgaag 1260  
caatagctga ttatttattt cttcttccaa tctatggta ctaacaacga caccgactaa 1320  
tgtttccgga gatagatttc caaagataca cacatttagga tacagactgt tataatcaaa 1380  
gattaataca ttattactaa acatctttt tttggagca aataacctac cgccttcata 1440  
agaaaactt tgtttggtt ctgatctaac taagatagtt ttagttcca acaatagctt 1500  
taacagtggc cccttgcgtga ctgtactcgc tctatattcg aataccatgg attgaggaag 1560  
cacatatgtt gacgcaccccg cgtctgttt tttttctact ccataatact cccacacaata 1620  
ctgacacacaaa caagcatcat gaatacagta tctagccata tctaaagcta ttttttagatt 1680  
ataatcctta tacatctgag ctaaatcaac gtcatccttt ccgaaagata atttatatgt 1740  
atcatttaggt aaagttaggac ataatagtagc gactttaat ccattttccc aaatatctt 1800

5 acgaattact ttacatataa tatecctcatc aacagtacaca taattacctg tggttaaaaac 1860  
 ctttgc当地 gtagccgtt tgc当地tcgc gtctgttagta tcgtcaccga taaacgtcat 1920  
 ttctctaact cctctatcca atactttacc catgcaactg aacgc当地tct tggatataga 1980  
 atccaatttg tacgaatcca atttttcaga ttttgaatg aatgaatata gatcgaaaaa 2040  
 tatagttcca ttattgttat taacgtgaaa cgtatgttgc gccatgccc ctactccctt 2100  
 atgactagac tgatttctct cataaaataca gagatgtaca gtttccctt tgc当地ggaga 2160  
 tcttaagata atcttctctc ctgttaataa ctctagacga tttagtaat atctcagatc 2220  
 aaagttatgt ccgttaaagg taacgacgtg gtc当地acgtt agttccaaca attgtttagc 2280  
 tattcgtaaac aaaactatcc cagaacatag aacttagttt cgttgc当地at ccatttccat 2340  
 10 tagtgc当地t atcctcaaacc atcctctatc gacggcttct ttttgc当地t gttccgtttaa 2400  
 catctctca ttaatgagcg taaacaataa tc当地tttacca ct当地atcga tataacagta 2460  
 acttgc当地tgc gagatgggt taataaatac agaaggaaac ttcttgc当地t agtgc当地ctc 2520  
 tatatctaga aataagtacg atcttgggat atcgaatcta ggtatcccc tagc当地aaaca 2580  
 15 gttacgtgga tc当地cacaat gataacatcc attttgc当地t ttgtcaat attgtc当地tgc 2640  
 caacgagtaa catccgtctg gagatatccc gtttagaaata taaaaccaac taatatttgc 2700  
 aatttgc当地tcc atggggcat ttgtatgtc gc当地tttctt ggcttctca tcaaccacat 2760  
 atctgc当地gacg gagcattttc tatctttaat atcttagatta taacttatttgc ttc当地tcaat 2820  
 gttctataggc ttc当地tcc ccaacggcc tgc当地ttaat ggaggaggag acaatgactg 2880  
 20 atatatttgc tccgtcaat cgtataaaaa gtaatgagga aatcgatataa atacggctctc 2940  
 accatttgc当地tccatc cgtataaaaa gtaatgagga aatcgatataa atacggctctc 3000  
 attaatgc当地tccatc cgtataaaaa gtaatgagga aatcgatataa atacggctctc 3021

25 <210> 8  
 <211> 1908  
 <212> DNA  
 <213> Vaccinia virus

30 <400> 8  
 ttataatttt accatctgac tc当地ggatttcc attaatatct ttataagagc tactaacgt 60  
 taattcttta taactgaact gagatataa caccggatct atggggccatc taatttgc当地t 120  
 aatgaatgtc cgccatataac taatggcaaa ttttgc当地tcaac aacgaaatc tactagatgt 180  
 gttaaaggta atatttcttca tgagcttgc当地t caataatcc ttgttgc当地t ctgc当地tccaa 240  
 35 gtc当地tccatc atcttgc当地tccatc tatccatgttcc accgttttta agttctggaa tatttgc当地t 300  
 ccattgtaaa gccc当地ttaat cgc当地tccatc atatctgtc ctgatagc当地t tttcaatatac 360  
 gacggacgtc当地tccatc aatactgtaa taaagggttgc当地t agtattgtc当地t tc当地tgc当地tccatc 420  
 aatatggtgc当地tccatc tttagtagtgc当地tccatc cggtaatcc acacaacgtc当地tccatc atatataact ttccctttgt 480  
 accatttttccatc acgtatgttgc当地tccatc gacgttgc当地tccatc aggttatttgc当地tccatc ttgttgc当地tccatc 540  
 aacagatttgc当地tccatc atacgatatt ttgttgc当地tccatc ctgatttgc当地tccatc actataatatac aatctagaca 600  
 40 gataatgtatc tccatc atacgatatt ttgttgc当地tccatc gagaaggatc atcgatgttgc当地tccatc ggataatatac tcccccatttcc 660  
 agtattctgc当地tccatc gataatgtatc taatgacact agttaaaggaaac atgttcttca ttcttagaaaaa 720  
 cggaaaatccatc ctacatggac tc当地tccatc ttcttgc当地tccatc cctgatgttgc当地tccatc tctc当地tccatc 780  
 ctctgtatc当地tccatc gatttcaagg atgccc当地tccatc ttcttgc当地tccatc aacgatccatc aatttgc当地tccatc 840  
 45 agcatctgtatc当地tccatc ttcttgc当地tccatc aatcgatatac tc当地tccatc ttcttgc当地tccatc cccatcgatatac 900  
 acaatagtc当地tccatc tgtaaaggata aaccgc当地tccatc aaatttatac gcatccatcc aaataaccctt 960  
 agcaccatcg gatgatatac atgttatttgc当地tccatc atagattttgc当地tccatc catccatc当地tccatc tatttggccatc 1020  
 gtataatgttgc当地tccatc agcaacggatc当地tccatc tatcgatatac attactctgc当地tccatc taacctacta gaatgatatac 1080  
 tc当地tccatc gtc当地tccatc ttcttgc当地tccatc atctatccatc ttcttgc当地tccatc gatcttgc当地tccatc 1140  
 actgttaaagg ttaacaaagg tattaccggatc当地tccatc gtacgttgc当地tccatc ttcttgc当地tccatc gatcttgc当地tccatc 1200  
 50 attatc当地tccatc atagctccatc aaactgatatac ggc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc 1260  
 cgaggaagaa attc当地tccatc acaatttgc当地tccatc tcttagatatac tatttgc当地tccatc cccatcgatatac 1320  
 cacatcaatt tc当地tccatc tcttagatatac tatttgc当地tccatc tatttgc当地tccatc aatgtatcgatatac 1380  
 tattatgc当地tccatc attatc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc gatcttgc当地tccatc 1440  
 55 ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc 1500  
 ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc 1560  
 ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc 1620  
 ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc 1680  
 ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc 1740  
 60 ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc 1800  
 ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc 1860  
 ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc ttcttgc当地tccatc 1908

5           <210> 9  
 <211> 1656  
 <212> DNA  
 <213> Vaccinia virus

10          <400> 9  
 ttagttatta tctccataa tcctggtaat acttaccct tgatcgtaag atacctata 60  
 caggtcatta catacaacta ccaattgttt ttgtacataa tagattggat gttgacatc 120  
 catggggaa taaactactc gaacagatag tttatcttc cccctagata cattggccgt 180  
 aatagtgtc ggcccaaaga atatcttgg tgtaaaggta aaagtttaggg ttcttgccc 240  
 attattgctt tttgtcagta gttcattata aattctcgag atgggtccgt tctctgaata 300  
 tagaacatca tttccaaatc taacttctag tctagaaata atatcggtct tattcttaa 360  
 15          atctattccc ttgtatgaagg gatcgtaat gaacaaatcc ttggcctttg attcggtctga 420  
 tctattatct ccgttataga cggtacgtt actagtccaa agacttacag gaatagatgt 480  
 atcgatgatg ttgatactat gtgatatgtg agcaaaagatt gttcttttag tggcatcact 540  
 atatgttcca gtaatggccgaaaactttt agaaatgtta tatataaaag aatttttcg 600  
 20          tgttccaaac attagcagat tagtatgaag ataaacactc atattatcg gaacattatc 660  
 aattttaca tacacatcg catcttgaat agaaaacgata ccatttctg gaacctctac 720  
 gatctcgcca gactccggat aaccagtccg tggccatca ctaacaataa ctagatcatc 780  
 caacaatcta ctcacatatg catctatata atcttttca tcttgaggt accctggata 840  
 cggaaaataat ttattatccg tatttccata ataagggtta gtataaacag agagcgatgt 900  
 25          tgccgcatga acttcagttt cagtcgccccgt tggttgggtt atttgaccta ttactctcct 960  
 agggttctct ataaacgatg gtttaattt tacattctta accatatatac caataaagct 1020  
 caattcagga acataaacaa attctttgtt gaacgttca aagtcaacg aagagtcacg 1080  
 aataacgata tcggatactg gattgaaggt taccgttacg gtaattttt aatcggtatag 1140  
 tttaagactg ctgaatgtat cttccacatc aaacggagtt ttaatataaa cgtataactgt 1200  
 agatggttct ttaatagtgtt cattaggagt taggccaata gaaatatcat taagttcact 1260  
 30          agaatatcca gaggtttca aagcaattgtt attattgata caattattat ataattctc 1320  
 gcccctcaatt tcccaaataa caccgttaca cgaagagata gatacgatgat taatacattt 1380  
 atatccaaca tatgtacgt aaccgaatct tcccatatc ttaatttctg gaagttccaa 1440  
 actcagaacc aaatgattaa gcgcagtaat atactgatcc ctaatttcga agctagcgat 1500  
 35          agcctgattt tctggaccat cgttgtcat aactccggat agagaaatatttgcggat 1560  
 atataaaggat ggaatttgc tatcgactgc gaagacatta gaccgtttaa tagagtcatc 1620  
 cccaccgatc aaagaattaa tgatagtatt attcat 1656

40          <210> 10  
 <211> 915  
 <212> DNA  
 <213> Vaccinia virus

45          <400> 10  
 ctagtttgtt ttttctcgcg aatatcgatc actcataaga aagagaatag cggttaagttat 60  
 aaacacgaat actatggcaa taattgcgaa tggtttattt cttcgatataatttgtata 120  
 atatgaaaaa catgtctctc tcaaatcgga caaccatctc ataaaatagt tctcgccgc 180  
 tggagaggta gttgtcttc gtataatctc cccagaataa tataacttgcg tgcgtcgtt 240  
 caatttatac ggatttctat agttctctgt tatataatac gttttccat catgattaga 300  
 50          cgacgacaat agtggctaa atttagatag ttgatcgaa tgaatgtta tggcggtgg 360  
 aaaaattatc catacagcgt ctgcagatgtt ctgtatagtt gttccttagat atgtaaaata 420  
 atccaacgta ctatgttagca aattgtcttag ataaaatact gaatcaaacg ggcgcagacgt 480  
 attagcgat ctaatggat ccaattgtatt gactatctt taaaatata cattttatg 540  
 atccgataact tgtaagaata tagaaataat gataagtcca tcattgtgtt ttttgcctc 600  
 55          ttcataagaa ctatatttt tcttattcca atgaacaaga ttaatctctc cagagtattt 660  
 gtacacatct atcaagtgtat tggatccata atcgatcttc ttcccccaat atatatgttag 720  
 tgatgataac acatattcat tggggagaaa ccctccactt atatatctc ctttaaaatt 780  
 aatccctact agtttccag tggatctggat agtgggttgcgatcactatataatgtat 840  
 gtctaacgcg ttcataatcgcg cgttagaaat tgcttttta gtttctatataataggaga 900  
 60          tagttgtgc ggcatt 915

&lt;210&gt; 11

&lt;211&gt; 939

&lt;212&gt; DNA

5 &lt;213&gt; Vaccinia virus

&lt;400&gt; 11

ttattcagca ttacttgata tagtaatatt aggcacagtc aaacattcaa ccactctcg 60  
 tacattaact ctctcatttt ctttaacaaa ttctgcaata tcttcgtaaa aagattctt 120  
 10 aaacctttta gaatatctat cgactctaga tgaatagcg ttctgtcaaca tactatgtt 180  
 tgtatacata aaggcgccca ttttaacagt ttcttagtgc aaaatgtcg cgatcctagg 240  
 atcctttaga atcacataga ttgacgattc gtctcttta gtaactctag taaaataatc 300  
 atacaatcta gtacgcgaaa taatattatc cttgacttga ggagatctaa acaatctagt 360  
 tttgagaaca tcgataagg tcatcggaat gacatacata ctatcttta tagaactctt 420  
 15 ttcatccagt tgaatggatt cgtccttaac caactgatta atgagatctt ctatcttatac 480  
 attttcaga tgatatgtat gtccattaaa gttaaattgt gtacgcctc tttttagtct 540  
 agcagccaat actttaacat cactaatatc gatatacataa ggagatgtt tatctatgtt 600  
 attaagaatt cgttttcga catccgtcaa aaccaattcc ttttgcctg tatcatccag 660  
 20 tttccatcc tttgtaaaaga aattatttc tactagacta ttaataagac tgataaggat 720  
 tcctccataa ttgcacaatc caaactttt cacaacta gactttacga gatctacagg 780  
 aatgcgtact tcaggttct tagcttgta tttttctt tgccggacatt ttcttgta 840  
 caactcatct accatttcat tgatttttagc agtgaataaa gcttcaatg cacgggcaact 900  
 gatactattt gaaacgagtt gatcttcaaa ttccggccat 939

25

&lt;210&gt; 12

&lt;211&gt; 429

&lt;212&gt; DNA

30 &lt;213&gt; Vaccinia virus

&lt;400&gt; 12

ttaattgcaa agatctataat aatcattata gcgttgactt atggactctg gaatcttaga 60  
 cgtatgtacag tcatctataa tcatggcata ttttaatacat ttttttatag catagttagtt 120  
 35 atctacgtatg ttagatattt ctctcaatga atcaatcaca taatctaatac taggtttatg 180  
 acataatagc attttcagca gttcaatgtt tctagattcg ttgatggcaa tggctataca 240  
 tttatcccg ttatttgatec taatgttgac atctgaaccg gattctagca gtaaaagatac 300  
 tagagattgt ttatttatac taacagcctt gtgaagaagt gtttctccctc gtttgcataat 360  
 catgttaatg tctttaagat aaggttaggca aatgtttata gtaactaagaa ttgggcaagc 420  
 ataagacat 429

40

&lt;210&gt; 13

&lt;211&gt; 1038

&lt;212&gt; DNA

45 &lt;213&gt; Vaccinia virus

&lt;400&gt; 13

atggatatct tcagggaaat cgcacatcttct atgaaaggag agaatgtatt catttctcca 60  
 50 gctgtcaatct cgtcagtatt gacaatactg tattatggag ctaatggatc cactgctgaa 120  
 cagctatcaa aatatgtaga aaaggaggag aacatggata aggttagcgc tcaaaaatatc 180  
 tcattcaat ccataaataa agtataatgg cgatattctg ccgtgtttaa agattcctt 240  
 ttgagaaaaa ttggcgataa gtttcaact gttgacttca ctgatttgcg cactatagat 300  
 gcaatcaaca agtgtgtaga tatctttact gagggaaaaa tcaatccact attggatgaa 360  
 ccattgtctc ctgatacctg tctcctagca attagtgccg tatactttaa agcaaaaatgg 420  
 55 ttgacgcccattc tcgaaaagga atttaccatg gattatccct tttacgtatc tccgacggaa 480  
 atggtagatg taagtatgtat gtctatgtac ggcaaggcat ttaatcacgc atctgtaaag 540  
 gaatcattcg gcaacttttc aatcatagaa ctggccatatg ttggagatac tagtatgtatg 600  
 gtcattcttc cagacaagat tgatggatta gaatccatag aacaaaatct aacagataca 660  
 aatttttaaga aatgggtgtaa ctctctggaa gctacgttta tcgatgttca cattcccaag 720  
 60 tttaaaggtaa caggctcgta taatctgtg gatactctag taaagtctagg actgacagag 780  
 gtgttcgggtt caactggaga ttatagcaat atgtgtatt cagatgtgag tgcacgct 840

atgatccaca aaacgtatat agatgtcaat gaagagtata cagaaggcgc tgcagcaact 900  
tgtgcactgg tgtcagactg tgcatcaaca attacaaatg agttctgtt agatcatccg 960  
ttcatctatg tgattaggca tggttatggaa aaaattcttt tcgttggtag atattgctct 1020  
ccgacaacta attgttaa 1038

5

<210> 14  
<211> 453  
<212> DNA

10 <213> Vaccinia virus

<400> 14  
ttaatccatg gactcataat ctctatacgg gattaacggg tttcttat acggggatga 60  
gtagttctct tctttaactt tatactttt actaatcata ttttagactga tttatggta 120  
15 atagtgtttg aagagctcg tctcatcata agaataaaatc aatatctctg tttttttttt 180  
atacagatgt attacagcct catatattac gtaatagaac gtgtcatcta ctttattaaac 240  
tttcaccgca tagtttttgg caaatacggg taatcctttg acctcgctcg tttccgacca 300  
atctgggcgt ataatgaatc taaaactttaa tttcttgtaa tcattcgaaa taatttttag 360  
20 tttgcattccg tagttatccc ctttatgtaa ctgtaaattt ctcaacgcga tatctccatt 420  
aataatgtat tcgaattcg tctgtataacc cat 453

<210> 15  
<211> 660  
<212> DNA  
25 <213> Vaccinia virus

<400> 15  
atggcgtatgt ttacgcaca cgctctcggt ggttacgacg agaatcttca tgcctttcct 60  
30 ggaatatcat cgactgttgc caatgtatgc agggaaatatt ctgttggtc agtttataat 120  
aacaagtatg acattgtaaa agacaaatatt atgtgggtt acagtcagg gaacaagaga 180  
tatattggag cactgctgcc tatgtttgag tgcaatgaat atctacaaat tggagatccg 240  
atccatgatc aagaaggaaa tcaaattctt atcatcacat atgcaccaa aaactactat 300  
35 gctctaagcg gaatcgggta cgagagtcta gacttgggtt tggaggagt agggattcat 360  
catcacgtac ttgaaacacgg aaacgctgtt tatggaaaag ttcaacatga ttattctact 420  
atcaaagaga aggccaaaga aatgagtaca cttagtccag gacctataat tgattaccac 480  
gtctggatag gagattgtat ctgtcaagtt actgtgtgg acgtacatgg aaaggaaatt 540  
atgagaatga gattaaaaaa gggtgcgtg cttccgatcc caaatctgtt aaaagttaaa 600  
40 cttggggaga atgatacaga aaatcttct tctactatataat cggcggcacc atcgaggtaa 660

<210> 16  
<211> 957  
<212> DNA  
45 <213> Vaccinia virus

<400> 16  
atgacgaccc taccaagtgc ggtatataaa aacgatttaa ttacagagtt ttcagaagat 60  
50 aattatccat ctaacaaaaa ttatgaaata actcttcgtc aaatgtctat tctaactcac 120  
gttaacaacg tggtagatag agaacataat gcccggtag tgcacatctcc agaggaaata 180  
tcctcacaaac ttaatgaaga tctatttcca gatgtatgtt ctccggccac tattatcgaa 240  
agagtacaaac ctcataactac tattattgac gatactccac ctcctacgtt tcgttagagag 300  
ttattgataat cggacaacgc tcaacaacga gaaaaaagat ttaatattac agtacgaaa 360  
55 aatgctgaag caataatgg atcttagatct atgatatctt ctatgcaac acaaacacca 420  
tccttggggat tagttatgtt taaagataaa agaattcaga tggtaggg tgaagtgggtt 480  
aatcttagaa atcaacgtc taataaaaaa tcacatgtata atttagataa ttttaccaga 540  
atactatttgc gtaagactcc gtataaaatca acagaagttt ataaagctat agccatcggtt 600  
aattatgcaat atttgaacgg gtctccctta tcagtcgagg acttggatgt ttgttcagag 660  
60 gatgaaatag atagaatcta taaaacgatt aaacaatatc acgaaagtag aaaaacgaaaa 720  
attatcgatca ctaacgtatc tattattgtc ataaacatca tcgagcaagc attgctaaaa 780  
ctcgattttt aagaaatcaa aggactgatgtt accgatataca cttcagaaat tatecgatgtt 840

```
gagatcgagg atgactgcga tgctgttagca tctaaactag gaatcggtaa cagtcgggtt 900  
cttaatattt tattgtttat actcaagata ttccgttaaac gaattaaaat tatttaa 957
```

5 <210> 17  
<211> 273  
<212> DNA  
<213> Vaccinia virus

10 <400> 17  
ttagttcatg gaaatatacg ttagtgggt atgaatgact ccgcttaactc tgggggtgc 60  
gcagtgcatt ccccacatag aataaattag catccgact gtgataataa taccaagtat 120  
aaacgccata atactcaata ctttccatgt acgagtgggta ctggtagact tactaaagtc 180  
aataaaggcg aagatacacg aaagaatcaa aagaatgatt ccagcgatta gcacgcccga 240  
aaaataattt ccaatcataa gcatcatgtc cat 273

15 <210> 18  
<211> 3861  
<212> DNA  
<213> Vaccinia virus

20 <400> 18  
atggctgtaa tctctaaggtaacgtatgt ctatatgatc aaaaagagat taatgctaca 60  
gatattatca ttagtcatgt taaaaatgac gacgatatcg gtaccgttaa agatggtaga 120  
ctagggtcta tggatggggc attatgttaa acttgtgggaa aacggaatt ggaatgtttc 180  
ggtaactggg gtaaagaatgg tatttataaa actcatatag ttaagcctga atttatttca 240  
gaaattattt gtttactgaa tcataatgtt attcaactcg gattattgcg ttacagagaa 300  
ccgtattccg acgatattaa cctaaaagag ttatcgggac acgctcttag gagattaaag 360  
gataaaaat tatccaagaa aaagtcatgt tggAACAGCG aatgtatgca accgtatcaa 420  
aaaattactt ttccaaagaa aaagggttgc ttctgtcaaca agttggatga tattaacgtt 480  
cctaattctc tcataatca aaagttattt tctattcatg aaaagtttg gccattattt 540  
gaaattcata aatatccagc taacttattt tatacagact actttcccat ccctccgctg 600  
attattagac cggctattag ttttggata gatagtatac ccaaagagac caatgaatta 660  
acttacttat taggtatgtat cgtaaagaaat tgtaacttgc atgctgatga acaggttattc 720  
cagaaggcgg taatagaata cgatgatattt aaaattattt ctaataacac ttccagttatc 780  
aattttatcat atattacatc cggcaaaaat aatatgatta aggttatat tgcgccccga 840  
cgaaaagatc agaccgttag atctgtattt ggtcccgatc catctatcac cgtaatgag 900  
gttaggaatgc ccgcataatataca tagaaataca ttacagaaa agatatttgc taatgcctt 960  
40 acagtggata aagttaaaca actattagcg tcaaaaccaag taaaattttt ctttaataaa 1020  
cgattaaacc aattaacaag aatacgc当地 ggaaagttt taaaaataa aatacattt 1080  
ttgcctgggtt attgggttaga agtagctgtt caagaatataa caagtattat ttttggaaaga 1140  
cagccgtctc tacatagata caacgtcatc gcttcataatc tcagagctac cgaaggagat 1200  
actatcaaaa tatctccgg aattgccaac tctcaaaatg ctgatttgcg cggagatgaa 1260  
gaatggatga tattggagca aaatcctaaa gcccgtatgt aacaaagtat tcttattgtat 1320  
ccgacgacgt tactcaaaaca cgatattcat ggagcccccg tttatggatc tattcaagat 1380  
gaaatcgtag cagcgatattt attgttttaga atacaagatc ttgttttaga tgaagtattt 1440  
aacatcttgg ggaaatatgg aagaaaggcc gatcctaaag gtaaatgtaa attcagcggt 1500  
aaagatatct atacttactt gataggtgaa aagattaattt atccgggtct cttaaaggat 1560  
ggtaaaatattt ttgc当地acgc cgtagatagt aattttgtt tggctatgag gcatctgtca 1620  
ttggctggac tcttataatca tcataatgtc aacgtggaaat gtagtcaactt tattatcaag 1680  
tcataatcttattt ttttaagag atatctatctt attacgggtt ttgggttgac attcaagat 1740  
ctgagaccaa attcgacgtt cactaataaa ttggaggccat tcaacgtaga aaaaatagaa 1800  
cttatcaaaag aagcatatcgc caaatatctc aacgtgttac gtagtcaactt gacaaatctt 1860  
ttatctaaag ctttagaggc ggactatgtt gatccatgtt aatatccgag agatagaaga acatatgaga 1920  
ctccctggaaa tggccaaagc gggttataaa gtaaatccca tagatgtcc agataataac 1980  
ggtaactttag gacaacagag gattgtatgtt gaaccagcggat cagaactaat gtatattcta 2040  
gtcttacattt actatcttcc agactctaa gatccagaag agactcgatg attgggttaga 2100  
tctttaacaa aaggatttaac aggttctcaa tattactttt cgtaccggaa cactggctag aaaaatcatt 2160  
caatctactg atatcgatcg tgaaacatca 2220  
25 <210> 19  
<211> 3861  
<212> DNA  
<213> Vaccinia virus

30 <400> 19  
atggctgtaa tctctaaggtaacgtatgtt ctatatgatc aaaaagagat taatgctaca 60  
gatattatca ttagtcatgtt taaaaatgac gacgatatcg gtaccgttaa agatggtaga 120  
ctagggtcta tggatggggc attatgtttt acttgtgggaa aacggaatt ggaatgtttc 180  
ggtaactggg gtaaagaatgg tatttataaa actcatatag ttaagcctga atttatttca 240  
gaaattattt gtttactgaa tcataatgtt attcaactcg gattattgcg ttacagagaa 300  
ccgtattccg acgatattaa cctaaaagag ttatcgggac acgctcttag gagattaaag 360  
gataaaaat tatccaagaa aaagtcatgtt tggAACAGCG aatgtatgca accgtatcaa 420  
aaaattactt ttccaaagaa aaagggttgc ttctgtcaaca agttggatga tattaacgtt 480  
cctaattctc tcataatca aaagttattt tctattcatg aaaagtttg gccattattt 540  
gaaattcata aatatccagc taacttattt tatacagact actttcccat ccctccgctg 600  
attattagac cggctattag ttttggata gatagtatac ccaaagagac caatgaatta 660  
acttacttat taggtatgtat cgtaaagaaat tgtaacttgc atgctgatga acaggttattc 720  
cagaaggcgg taatagaata cgatgatattt aaaattattt ctaataacac ttccagttatc 780  
aattttatcat atattacatc cggcaaaaat aatatgatta aggttatat tgcgccccga 840  
cgaaaagatc agaccgttag atctgtattt ggtcccgatc catctatcac cgtaatgag 900  
gttaggaatgc ccgcataatataca tagaaataca ttacagaaa agatatttgc taatgcctt 960  
40 acagtggata aagttaaaca actattagcg tcaaaaccaag taaaattttt ctttaataaa 1020  
cgattaaacc aattaacaag aatacgc当地 ggaaagttt taaaaataa aatacattt 1080  
ttgcctgggtt attgggttaga agtagctgtt caagaatataa caagtattat ttttggaaaga 1140  
cagccgtctc tacatagata caacgtcatc gcttcataatc tcagagctac cgaaggagat 1200  
actatcaaaa tatctccgg aattgccaac tctcaaaatg ctgatttgcg cggagatgaa 1260  
gaatggatga tattggagca aaatcctaaa gcccgtatgt aacaaagtat tcttattgtat 1320  
ccgacgacgt tactcaaaaca cgatattcat ggagcccccg tttatggatc tattcaagat 1380  
gaaatcgtag cagcgatattt attgttttaga atacaagatc ttgttttaga tgaagtattt 1440  
aacatcttgg ggaaatatgg aagaaaggcc gatcctaaag gtaaatgtaa attcagcggt 1500  
aaagatatct atacttactt gataggtgaa aagattaattt atccgggtct cttaaaggat 1560  
ggtaaaatattt ttgc当地acgc cgtagatagt aattttgtt tggctatgag gcatctgtca 1620  
ttggctggac tcttataatca tcataatgtc aacgtggaaat gtagtcaactt tattatcaag 1680  
tcataatcttattt ttttaagag atatctatctt attacgggtt ttgggttgac attcaagat 1740  
ctgagaccaa attcgacgtt cactaataaa ttggaggccat tcaacgtaga aaaaatagaa 1800  
cttatcaaaag aagcatatcgc caaatatctc aacgtgttac gtagtcaactt gacaaatctt 1860  
ttatctaaag ctttagaggc ggactatgtt gatccatgtt aatatccgag agatagaaga acatatgaga 1920  
ctccctggaaa tggccaaagc gggttataaa gtaaatccca tagatgtcc agataataac 1980  
ggtaactttag gacaacagag gattgtatgtt gaaccagcggat cagaactaat gtatattcta 2040  
gtcttacattt actatcttcc agactctaa gatccagaag agactcgatg attgggttaga 2100  
tctttaacaa aaggatttaac aggttctcaa tattactttt cgtaccggaa cactggctag aaaaatcatt 2160  
caatctactg atatcgatcg tgaaacatca 2220  
55 <210> 20  
<211> 3861  
<212> DNA  
<213> Vaccinia virus

60 <400> 20  
atggctgtaa tctctaaggtaacgtatgtt ctatatgatc aaaaagagat taatgctaca 60  
gatattatca ttagtcatgtt taaaaatgac gacgatatcg gtaccgttaa agatggtaga 120  
ctagggtcta tggatggggc attatgtttt acttgtgggaa aacggaatt ggaatgtttc 180  
ggtaactggg gtaaagaatgg tatttataaa actcatatag ttaagcctga atttatttca 240  
gaaattattt gtttactgaa tcataatgtt attcaactcg gattattgcg ttacagagaa 300  
ccgtattccg acgatattaa cctaaaagag ttatcgggac acgctcttag gagattaaag 360  
gataaaaat tatccaagaa aaagtcatgtt tggAACAGCG aatgtatgca accgtatcaa 420  
aaaattactt ttccaaagaa aaagggttgc ttctgtcaaca agttggatga tattaacgtt 480  
cctaattctc tcataatca aaagttattt tctattcatg aaaagtttg gccattattt 540  
gaaattcata aatatccagc taacttattt tatacagact actttcccat ccctccgctg 600  
attattagac cggctattag ttttggata gatagtatac ccaaagagac caatgaatta 660  
acttacttat taggtatgtat cgtaaagaaat tgtaacttgc atgctgatga acaggttattc 720  
cagaaggcgg taatagaata cgatgatattt aaaattattt ctaataacac ttccagttatc 780  
aattttatcat atattacatc cggcaaaaat aatatgatta aggttatat tgcgccccga 840  
cgaaaagatc agaccgttag atctgtattt ggtcccgatc catctatcac cgtaatgag 900  
gttaggaatgc ccgcataatataca tagaaataca ttacagaaa agatatttgc taatgcctt 960  
40 acagtggata aagttaaaca actattagcg tcaaaaccaag taaaattttt ctttaataaa 1020  
cgattaaacc aattaacaag aatacgc当地 ggaaagttt taaaaataa aatacattt 1080  
ttgcctgggtt attgggttaga agtagctgtt caagaatataa caagtattat ttttggaaaga 1140  
cagccgtctc tacatagata caacgtcatc gcttcataatc tcagagctac cgaaggagat 1200  
actatcaaaa tatctccgg aattgccaac tctcaaaatg ctgatttgcg cggagatgaa 1260  
gaatggatga tattggagca aaatcctaaa gcccgtatgt aacaaagtat tcttattgtat 1320  
ccgacgacgt tactcaaaaca cgatattcat ggagcccccg tttatggatc tattcaagat 1380  
gaaatcgtag cagcgatattt attgttttaga atacaagatc ttgttttaga tgaagtattt 1440  
aacatcttgg ggaaatatgg aagaaaggcc gatcctaaag gtaaatgtaa attcagcggt 1500  
aaagatatct atacttactt gataggtgaa aagattaattt atccgggtct cttaaaggat 1560  
ggtaaaatattt ttgc当地acgc cgtagatagt aattttgtt tggctatgag gcatctgtca 1620  
ttggctggac tcttataatca tcataatgtc aacgtggaaat gtagtcaactt tattatcaag 1680  
tcataatcttattt ttttaagag atatctatctt attacgggtt ttgggttgac attcaagat 1740  
ctgagaccaa attcgacgtt cactaataaa ttggaggccat tcaacgtaga aaaaatagaa 1800  
cttatcaaaag aagcatatcgc caaatatctc aacgtgttac gtagtcaactt gacaaatctt 1860  
ttatctaaag ctttagaggc ggactatgtt gatccatgtt aatatccgag agatagaaga acatatgaga 1920  
ctccctggaaa tggccaaagc gggttataaa gtaaatccca tagatgtcc agataataac 1980  
ggtaactttag gacaacagag gattgtatgtt gaaccagcggat cagaactaat gtatattcta 2040  
gtcttacattt actatcttcc agactctaa gatccagaag agactcgatg attgggttaga 2100  
tctttaacaa aaggatttaac aggttctcaa tattactttt cgtaccggaa cactggctag aaaaatcatt 2160  
caatctactg atatcgatcg tgaaacatca 2220  
55 <210> 21  
<211> 3861  
<212> DNA  
<213> Vaccinia virus

	aaaaagatgg	aggatatgg	ggtcgacgga	tacggacaag	tagttatagg	taatacgctc	2340
	atcaagtacg	ccgccaatta	tacccaaaatt	ctaggctca	tatgtaaacc	tgttagatctt	2400
	atcttatccag	atgagtccat	gacttggtat	ttggaaatta	gtgctctgt	gaataaaata	2460
5	aaacaggat	tcgttactc	tcagaacacag	aaacttgcaa	aaaagacatt	ggcgccgtt	2520
	aatttcctag	tattcgtcaa	acccaccact	gaggataatg	ctattaaggt	taaggatctg	2580
	tacgatatga	ttcataaacgt	cattgatgat	gtgagagaga	aataacttctt	tacggtatct	2640
	aatatagatt	ttatggagta	tatattctt	acgcacatcta	atccttctag	aattagaatt	2700
10	acaaaagaaa	cggctatcac	tatcttgaa	aatgtctatg	aaaaactcaa	ttatactcta	2760
	gttggtgaaa	ctccatttgg	aattatttct	gcacaggat	tgtctgagaa	gtttacacaa	2820
	caaggccctgt	ccagtttca	cactactgaa	aaaagtggtg	ccgtcaaaaca	aaaacttggt	2880
	ttcaacgagt	ttaataacct	gactaattt	agtaagaata	agaccgaaat	tatcactctg	2940
	gtatccgatg	atatctctaa	acttcaatct	gttaagatta	atttcaattt	tgtatgtttg	3000
	ggagaattaa	atccaaacat	cactcttcga	aaagaaacag	ataggatgt	agtagatata	3060
	atagtcaata	gattatacat	caagagagca	gaaatttaccc	aatttagtcgt	cgaatataatg	3120
15	attgaacgat	ttatctcttt	tagcgtcatt	gtaaaaggaaat	ggggatgttgc	gacattcatt	3180
	gaggacgagg	ataatattag	atttactgtc	tacctaaatt	tcgttgaacc	ggaagaatttgc	3240
	aatcttagta	agtttatgtat	ggttcttccg	ggtgccgc	acaaggggcaa	gatttagtaaa	3300
	ttcaagatcc	ctatctctga	ctatacggga	tatgacgact	tcaatcaaac	aaaaaaagctc	3360
20	aataagatga	ctgtagaact	catgaatcta	aaagaattgg	gttcttcga	tttggaaaac	3420
	gtcaacgtgt	atccctggagt	atggaatata	tacgatatct	tcggatctcg	ggccgctcg	3480
	gaataacttgt	gcgaagccat	gttaaaccacc	tatggagaag	gttccgatta	tctgtatctcg	3540
	ccttgtgtatc	ttctcgctag	tttactatgt	gctagttacg	aaccagaatc	agtgaataaa	3600
	ttcaagttcg	gcgcagctag	tactcttaag	agagctacgt	tcggagacaa	taaagcatttgc	3660
25	ttaaacgcgg	ctcttcataaa	aaagtcaagaa	cctttaacg	ataatagtag	ctgccactt	3720
	tttagcaagg	tcccttaatat	aggaacttgg	tattacaat	actttatcg	cttgggtctt	3780
	ctcatgagaa	tggaaaggaa	actatctgtat	aagatatctt	ctcaaaaagat	caaggaaatg	3840
	gaagaaacacag	aagactttta	a				3861

30 <210> 19  
<211> 2001  
<212> DNA  
<213> Vaccinia virus



ttttctctac aacgacgatt ggaggattta gtcctgata agttatggga accagggtt 720  
 attcattcg aagacgctat aaaaagagtt tc当地atc tataataaac 780  
 ttaatgatc tc当地aaaa taatttaca acggtaccac tggtcataga ttacgtaaca 840  
 ccttgc当地 tatgaaaaaa acgatcgcat aaacatccgc atcaactatc gttggaaaat 900  
 5 ggtgctatta gaatttacaa aactggtaat ccacatagt gt当地agttaa attgttccg 960  
 ttggatggta ataaactgtt taatattgca caaagaattt tagacactaa ctctgttta 1020  
 ttaaccgaac gaggagacta tatagttgg attaataatt catggaaatt taacagcgaa 1080  
 gaacccttga taacaaaact aattctgtca ataagacatc aactaccaa ggaatattca 1140  
 agcgaattac tctgtccgag gaaacgaaag actgttagaag ctaacatc agacatgtt 1200  
 10 gtagattcag tagagaccga tacctatccg gataaacttc cg当地aaaaa tggtgtattt 1260  
 gacctggtag acggaatgtt ttactctgga gatgatgcta aaaaatatac gtgtactgt 1320  
 tcaaccggat ttaaatttga cgatacaaag ttctgtcaag acagttccaga aatggaagag 1380  
 ttaatgaata tc当地aacga tatccaacca ttaacggatg aaaataagaa aaatagagag 1440  
 ctatatgaaa aaacattatc tagttgttta tgggtgcta ccaaaggatg ttaacattc 1500  
 15 tttttggag aaactgcaac tggaaagtgc acaacccaaac gtttggtaaa gtctgctatc 1560  
 ggtgacctgt ttgttgagac gggtaaaca attttacag atgtatttggta taaaggacct 1620  
 aatccatttta tc当地acat gc当地ggaaa agatctgtat tctgttagcga actacctgt 1680  
 tttgcctgta gtggatcaaa gaaaatttga tctgataata tt当地aaagtt gacagaacct 1740  
 tgggtcatgt gaagaccgtg ttctccaat aaaattaata atagaaacca tgctacaatc 1800  
 20 attatcgata ctaattacaa acctgtctt gataggatag ataacccatt aatgagaaga 1860  
 attgccgtcg tgc当地ttcag aacacactt tctcaaccc tt当地tagaga ggc当地ctgaa 1920  
 aataatgacg cgtacgataa agtcaaacta tt当地acgagg gtttagatgg taaaatacaa 1980  
 aataatagat atagatttgc atttctatac ttgttggta aatggtacaa aaaatatcat 2040  
 gttcctatta tgaaactata tc当地cacccc gaagagattc ct当地ttgc attctatctc 2100  
 25 aaaataggtt ctctgttagt atctagctt gtaaaggatc ttccatttaat gacggaccc 2160  
 tccaaaagg gatataattt gtacgataat gtgggtactc ttccgttgac tactttccaa 2220  
 cagaaaatattt ccaagttttt taattctaga ctatggac acgatataga gagcttcatc 2280  
 aatagacata agaaatttgc caatgttagt gatgaatatac tgcaatataat attcatagag 2340  
 gatatttcat ctccgtaa 2358  
 30

<210> 22  
 <211> 612  
 <212> DNA  
 35 <213> Vaccinia virus

<400> 22  
 ttaataatcg tc当地tattta aactgtttaa tgggtgata tcaacatcta ctttatttcc 60  
 40 cgc当地tataa ggttggtaa aggtataactg tt当地ggaaatg gttacattta tacttcttct 120  
 atagtcctgt ct当地cgatgt tc当地cacata tgcaaaagaac agaataaaaca aaataatgt 180  
 agaaaataataa tt当地atatct gt当地attcgta aataacattt attgc当地taa taatttacagc 240  
 agctacaata cacacaatag acatcccac agtggccca tt当地ccac gatacattt 300  
 agttaactaag caataggtaa taactaagct agtaagaggc aatggaaaag atgagataaa 360  
 tatcatcaat atagagatta gaggaggct atatagagcc aagacgaaca aaatcaaacc 420  
 45 gagtaaacgtt ctaacatcat tattttgaa gattcccaa taatcattca tt当地ccata 480  
 atcgttttgc atc当地acctc catctttagg cataaaacgt tgctgtgtt cctctgtaaa 540  
 taaatcttta tcaagcactc cagcaccgc agagaagtgc tcaagcatat tgtaatatct 600  
 taaataactc at 612

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
6. September 2002 (06.09.2002)

PCT

(10) Internationale Veröffentlichungsnummer  
**WO 02/068682 A3**

(51) Internationale Patentklassifikation<sup>7</sup>: C12Q 1/68,  
C12N 15/10, 15/39, C07K 14/07

Klinik, Obere Zahlbacher Strasse 63, 55131 Mainz  
(DE). **TUERECI, Özlem** [DE/DE]; III Medizinische  
Klinik, Obere Zahlbacher Strasse 63, 55131 Mainz (DE).  
**LUDEWIG, Burkhard** [DE/DE]; III Medizinische  
Klinik, Obere Zahlbacher Strasse 63, 55131 Mainz (DE).

(21) Internationales Aktenzeichen: PCT/EP02/01909

(74) Anwälte: **VOSSIUS, Volker** usw.; Holbeinstrasse 5,  
81679 München (DE).

(22) Internationales Anmeldedatum:  
22. Februar 2002 (22.02.2002)

(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT,  
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR,  
CU, CZ, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,

(25) Einreichungssprache: Deutsch

[Fortsetzung auf der nächsten Seite]

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:  
101 08 626.1 22. Februar 2001 (22.02.2001) DE

(71) Anmelder und

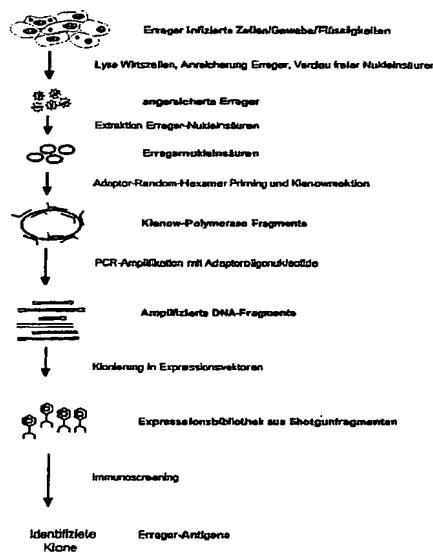
(72) Erfinder: **SAHIN, Ugur** [TR/DE]; III Medizinische

(54) Title: METHOD FOR IDENTIFYING BIOLOGICALLY ACTIVE STRUCTURES OF MICROBIAL PATHOGENS

(54) Bezeichnung: VERFAHREN ZUR IDENTIFIZIERUNG BIOLOGISCH AKTIVER STRUKTUREN MIKROBIELLER ERREGER

**(57) Abstract:** The invention relates to a method for identifying biologically active structures which are coded by the genome by microbial pathogens, starting from genomic controlled nucleic acids.

**(57) Zusammenfassung:** Die vorliegende Erfindung betrifft ein Verfahren zur Identifizierung biologisch aktiver Strukturen, die durch das Genom von mikrobiellen Erregern kodiert werden, ausgehend von genetischen Erreger-nukleinsäuren.



WO 02/068682 A3



SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,  
VN, YU, ZA, ZM, ZW.

**Veröffentlicht:**

— mit internationalem Recherchenbericht

**(84) Bestimmungsstaaten (regional):** ARIPO-Patent (GH,  
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ,  
TM), europäisches Patent (AT, BE, CH, CY, DE, DK,  
ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR),  
OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,  
ML, MR, NE, SN, TD, TG).

**(88) Veröffentlichungsdatum des internationalen  
Recherchenberichts:**

12. September 2003

Zur Erklärung der Zweibuchstaben-Codes und der anderen  
Abkürzungen wird auf die Erklärungen ("Guidance Notes on  
Codes and Abbreviations") am Anfang jeder regulären Ausgabe  
der PCT-Gazette verwiesen.

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP 02/01909

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 7 C12Q1/68 C12N15/10 C12N15/39 C07K14/07

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 C12Q C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, WPI Data, PAJ, MEDLINE, SEQUENCE SEARCH

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LANG SUSAN ET AL: "Identification of a novel antigen from <i>Staphylococcus epidermidis</i> ." FEMS IMMUNOLOGY AND MEDICAL MICROBIOLOGY, vol. 29, no. 3, November 2000 (2000-11), pages 213-220, XP002225330 ISSN: 0928-8244 the whole document	1,2,4,7, 8,20-24
Y	---	3,5,6, 9-19,25

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

## ° Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search  17 December 2002	Date of mailing of the international search report  02.05.03
Name and mailing address of the ISA  European Patent Office, P.B. 5818 Patentlaan 2 .NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer  Seranski, P

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/01909

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PADMALAYAM INDIRA ET AL: "Molecular cloning, sequencing, expression, and characterization of an immunogenic 43-kilodalton lipoprotein of Bartonella bacilliformis that has homology to NlpD/LppB." INFECTION AND IMMUNITY, vol. 68, no. 9, September 2000 (2000-09), pages 4972-4979, XP002225331 ISSN: 0019-9567 the whole document	1,2,4,7, 8,20-24
Y	---	3,5,6, 9-19,25
Y	WO 90 09457 A (BIOPHARM INC) 23 August 1990 (1990-08-23) page 5 -page 6	10-19,25
Y	US 5 731 171 A (BOHLANDER STEFAN K) 24 March 1998 (1998-03-24) column 2, line 55 -column 9, line 31	10,11
	-----	

**INTERNATIONAL SEARCH REPORT**

International application No.

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3.  Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-25

**Remark on Protest**

The additional search fees were accompanied by the applicant's protest.



No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/EP 02/01909

The International Searching Authority has determined that this international application contains multiple (groups of) inventions, namely

**Invention I: Claims 1-25**

Method for the identification of biological structures coded by the genome of microbial pathogens.

**Invention II: Claims 26-29 (in part)**

Vaccinia virus antigen coded by the nucleic acid with the SEQ ID NO. 4.

**Inventions III-XX: Claims 26-29 (in part in each case)**

Vaccinia virus antigens coded by the nucleic acids with the SEQ ID NO. 5-22.

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International Application No

PCT/EP 02/01909

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9009457	A 23-08-1990	AU	5404890 A	05-09-1990
		CA	2046919 A1	15-08-1990
		EP	0458909 A1	04-12-1991
		WO	9009457 A2	23-08-1990
US 5731171	A 24-03-1998	NONE		

## INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP 02/01909

A. KLASIFIZIERUNG DES ANMELDUNGSGEGENSTANDES  
IPK 7 C12Q1/68 C12N15/10

C12N15/39 C07K14/07

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

## B. RECHERCHIERTE GEBIETE

Recherchierte Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)

IPK 7 C12Q C12N

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

EPO-Internal, BIOSIS, WPI Data, PAJ, MEDLINE, SEQUENCE SEARCH

## C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie°	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	LANG SUSAN ET AL: "Identification of a novel antigen from <i>Staphylococcus epidermidis</i> ." FEMS IMMUNOLOGY AND MEDICAL MICROBIOLOGY, Bd. 29, Nr. 3, November 2000 (2000-11), Seiten 213-220, XP002225330 ISSN: 0928-8244 das ganze Dokument	1,2,4,7, 8,20-24
Y	---	3,5,6, 9-19,25
	-/-	



Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen



Siehe Anhang Patentfamilie

- ° Besondere Kategorien von angegebenen Veröffentlichungen :  
 "A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist  
 "E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist  
 "L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)  
 "O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht  
 "P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist
- "T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist  
 "X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erforderlicher Tätigkeit beruhend betrachtet werden  
 "Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erforderlicher Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist  
 "&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche	Absendedatum des internationalen Recherchenberichts
17. Dezember 2002	02.05.03

Name und Postanschrift der Internationalen Recherchenbehörde  
Europäisches Patentamt, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Seranski, P

## INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP 02/01909

## C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	PADMALAYAM INDIRA ET AL: "Molecular cloning, sequencing, expression, and characterization of an immunogenic 43-kilodalton lipoprotein of <i>Bartonella bacilliformis</i> that has homology to NlpD/LppB." INFECTION AND IMMUNITY, Bd. 68, Nr. 9, September 2000 (2000-09), Seiten 4972-4979, XP002225331 ISSN: 0019-9567 das ganze Dokument	1,2,4,7, 8,20-24
Y	---	3,5,6, 9-19,25
Y	WO 90 09457 A (BIOPHARMA INC) 23. August 1990 (1990-08-23) Seite 5 -Seite 6	10-19,25
Y	US 5 731 171 A (BOHLANDER STEFAN K) 24. März 1998 (1998-03-24) Spalte 2, Zeile 55 -Spalte 9, Zeile 31 -----	10,11

# INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen  
PCT/EP 02/01909

## Feld I Bemerkungen zu den Ansprüchen, die sich als nicht recherchierbar erwiesen haben (Fortsetzung von Punkt 2 auf Blatt 1)

Gemäß Artikel 17(2)a) wurde aus folgenden Gründen für bestimmte Ansprüche kein Recherchenbericht erstellt:

1.  Ansprüche Nr.  
weil sie sich auf Gegenstände beziehen, zu deren Recherche die Behörde nicht verpflichtet ist, nämlich
  
2.  Ansprüche Nr.  
weil sie sich auf Teile der internationalen Anmeldung beziehen, die den vorgeschriebenen Anforderungen so wenig entsprechen, daß eine sinnvolle internationale Recherche nicht durchgeführt werden kann, nämlich
  
3.  Ansprüche Nr.  
weil es sich dabei um abhängige Ansprüche handelt, die nicht entsprechend Satz 2 und 3 der Regel 6.4 a) abgefaßt sind.

## Feld II Bemerkungen bei mangelnder Einheitlichkeit der Erfindung (Fortsetzung von Punkt 3 auf Blatt 1)

Die internationale Recherchenbehörde hat festgestellt, daß diese internationale Anmeldung mehrere Erfindungen enthält:

siehe Zusatzblatt

1.  Da der Anmelder alle erforderlichen zusätzlichen Recherchengebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht auf alle recherchierbaren Ansprüche.
  
2.  Da für alle recherchierbaren Ansprüche die Recherche ohne einen Arbeitsaufwand durchgeführt werden konnte, der eine zusätzliche Recherchengebühr gerechtfertigt hätte, hat die Behörde nicht zur Zahlung einer solchen Gebühr aufgefordert.
  
3.  Da der Anmelder nur einige der erforderlichen zusätzlichen Recherchengebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht nur auf die Ansprüche, für die Gebühren entrichtet worden sind, nämlich auf die Ansprüche Nr.
  
4.  Der Anmelder hat die erforderlichen zusätzlichen Recherchengebühren nicht rechtzeitig entrichtet. Der internationale Recherchenbericht beschränkt sich daher auf die in den Ansprüchen zuerst erwähnte Erfindung; diese ist in folgenden Ansprüchen erfaßt:  
1-25

Bemerkungen hinsichtlich eines Widerspruchs

- Die zusätzlichen Gebühren wurden vom Anmelder unter Widerspruch gezahlt.  
 Die Zahlung zusätzlicher Recherchengebühren erfolgte ohne Widerspruch.

WEITERE ANGABEN

PCT/ISA/ 210

Die internationale Recherchenbehörde hat festgestellt, daß diese internationale Anmeldung mehrere (Gruppen von) Erfindungen enthält, nämlich:

Erfahrung I: Ansprüche 1-25

Verfahren zur Identifizierung von durch das Genom mikrobieller Erreger kodierter biologischer Strukturen

Erfahrung II: Ansprüche 26-29(teiweise)

Vacciniaivirusantigen kodiert durch die Nukleinsäure mit der Seq ID Nr 4

Erfahrung III-XX: Ansprüche 26-29(jeweils teiweise)

Vacciniaivirusantigene kodiert durch die Nukleinsäuren mit der Seq ID Nr 5-22

**INTERNATIONALES RECHERCHENBERICHT**

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP 02/01909

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie		Datum der Veröffentlichung
WO 9009457	A 23-08-1990	AU	5404890 A	05-09-1990
		CA	2046919 A1	15-08-1990
		EP	0458909 A1	04-12-1991
		WO	9009457 A2	23-08-1990
US 5731171	A 24-03-1998	KEINE		

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**